

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО**

Факультет інформатики та обчислювальної техніки

(назва факультету, інституту)

Кафедра автоматизованих систем обробки інформації і управління

(назва кафедри)

"На правах рукопису"

УДК 004.021

«До захисту допущено»

В.о.завідувача кафедри

О.А.Павлов

(підпис)

(ініціали, прізвище)

“ ” 20 19 р.

МАГІСТЕРСЬКА ДИСЕРТАЦІЯ

на здобуття ступеня магістра

за спеціальністю 126 Інформаційні системи та технології

(код та назва спеціальності)

ОПП Інформаційні управляючі системи та технології

(код та назва спеціалізації)

на тему: СИСТЕМА РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ

Виконав: студент VI курсу групи ІС-381мп

(шифр групи)

Терещенко Руслан Володимирович

(прізвище, ім'я, по батькові)

(підпис)

Науковий керівник доц., к.т.н, доц. Жданова О.Г.

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Консультант проф., д.т.н.Томашевський В.М.

(науковий ступінь, вчене звання, прізвище, ініціали)

(підпис)

Рецензент доц., к.т.н, доц. Рєпнікова Н.Б.

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Засвідчую, що у цій магістерській дисертації немає запозичень з праць інших авторів без відповідних посилань.

Студент

(підпис)

Київ – 2019

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ім. ІГОРЯ СІКОРСЬКОГО”

Факультет	<u>інформатики та обчислювальної техніки</u> (повна назва)
Кафедра	<u>автоматизованих систем обробки інформації та управління</u> (повна назва)
Рівень вищої освіти	<u>другий (магістерський) за освітньо-професійною програмою</u>
Спеціальність	<u>126 Інформаційні системи та технології</u> (код і назва)
ОПП	<u>Інформаційні управляючі системи та технології</u> (код і назва)

ЗАТВЕРДЖУЮ
В.о. завідувача кафедри
О.А.Павлов
(підпис) (ініціали, прізвище)
«___» _____ 2019 р.

ЗАВДАННЯ
на магістерську дисертацію студенту
Терещенко Руслану Володимирович
(прізвище, ім'я, по батькові)

1. Тема дисертації Система реєстрації та підтвердження авторських прав
науковий керівник дисертації доц., к.т.н, доц. Жданова Олена Григорівна
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)
- затверджені наказом по університету від “ 5 ” 11 20 19 р. № 3836-с
2. Строк подання студентом дисертації “ 2 ” 12 20 19 р.
3. Об'єкт дослідження
процес захисту авторських прав в Україні
4. Перелік завдань, які потрібно розробити
- дослідити способи збереження обмежених за обсягом даних про твір;
 - дослідити методи виявлення подібних за змістом документів;
 - розробити метод виведення цифрових відбитків творів, які дадуть змогу виявляти дублікати та близькі за змістом твори;
 - розробити інформаційну систему реєстрації та підтвердження авторських прав.

5. Орієнтовний перелік ілюстративного матеріалу

Діаграми способів використання системи для акторів

Діаграми процесів реєстрації, передачі та підтвердження прав на твір

Діаграма інформаційної моделі системи реєстрації та підтвердження авторських прав

Діаграма класів інформаційної системи реєстрації та підтвердження авторських прав

Діаграма розгортання системи реєстрації та підтвердження авторських прав

6. Орієнтовний перелік публікацій

Тези доповіді на міжнародній науково-практичній конференції

Тези доповіді на всеукраїнській науково-практичній конференції молодих вчених та студентів

7. Консультанти розділів дисертації

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

8. Дата видачі завдання “ 2 ” вересня 20 19 р.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Строк виконання етапів магістерської дисертації	Примітка
1	<i>Дослідження способів збереження обмежених за обсягом даних про твір</i>	<i>16.09</i>	
2	<i>Дослідження методів виявлення подібних за змістом документів</i>	<i>30.09</i>	
3	<i>Систематизація результатів огляду літератури</i>	<i>07.10</i>	
4	<i>Розробка методу виведення цифрових відбитків творів, які дадуть змогу виявляти дублікати та близькі за змістом твори</i>	<i>14.10</i>	
5	<i>Розробка інформаційної системи реєстрації та підтвердження авторських прав</i>	<i>28.10</i>	
7	<i>Проведення експериментальних досліджень розробленої системи</i>	<i>11.11</i>	
8	<i>Оформлення документації</i>	<i>18.11</i>	
9	<i>Подання роботи на попередній захист</i>	<i>20.11</i>	
10	<i>Подання роботи на основний захист</i>	<i>02.12</i>	

Студент

Р.В.Терещенко

Науковий керівник

О.Г.Жданова

РЕФЕРАТ

Магістерська дисертація: 105 с., 9 рис., 22 табл., 37 джерел, 1 додаток.

Актуальність. На даний час в Україні процес доведення авторського права, особливо якщо майнові права на твір передавались багато разів від одного власника до іншого, являється непростим та тривалим. Для цього потрібно надавати суду всі договори про передання авторських прав. Суд має ретельно дослідити кожний договір окремо, зокрема, конкретний обсяг прав, що передається за договором, та встановити, чи дійсно перейшли виключні майнові права від автора до нинішнього власника. Окрім того, передача прав можлива декільком власникам, а також часткова. Враховуючи нескінченну кількість усіх можливих комбінацій, доводити право власності на твір лише за допомогою людських ресурсів чи примітивних баз даних стає все складніше. Для вирішення цієї проблеми актуальною буде розробка спеціальної інформаційної системи, що автоматизує усі процеси реєстрації, передачі та доведення авторських прав, а також забезпечить надійний рівень захисту прав на твори своїх користувачів.

Мета дослідження – підвищення ступеню захисту авторських прав і зменшення вартості та часових рамок підтвердження авторства чи прав власності на твір.

Для досягнення мети необхідно виконати наступні **завдання**:

- дослідити способи збереження обмежених за обсягом даних про твір;
- дослідити методи виявлення подібних за змістом документів;
- розробити метод виведення цифрових відбитків творів, які дадуть змогу виявляти дублікати та близькі за змістом твори;
- розробити інформаційну систему реєстрації та підтвердження авторських прав.

Об'єкт дослідження – процес захисту авторських прав в Україні.

Предмет дослідження – інформаційна система реєстрації та підтвердження авторських прав.

Наукова новизна отриманих результатів. Розроблено метод виведення цифрових відбитків творів, які дають змогу виявляти не тільки дублікати, але й близькі за змістом твори шляхом поєднання алгоритмів MinHash та Кессак. Відбитки такого типу мають обмежену довжину і потребуватимуть обмеженого об'єму пам'яті для їх зберігання у інформаційній системі.

Публікації. Результати досліджень опубліковані у тезах чотирнадцятої міжнародної науково-практичної конференції “Математичне та імітаційне моделювання систем. МОДС '2019”, у тезах III всеукраїнської науково-практичної конференції молодих вчених та студентів “Інформаційні системи та технології управління” (ІСТУ-2019).

СИСТЕМА, АВТОРСЬКЕ ПРАВО, ФУНКЦІЯ ХЕШУВАННЯ, АЛГОРИТМ ХЕШУВАННЯ, ПОДІБНІ ДОКУМЕНТИ, ЦИФРОВИЙ ВІДБИТОК

ABSTRACT

Master's Thesis: 105 pages, 9 figures, 22 tables, 37 sources, 1 appendix.

Topicality. At present, in Ukraine, the process of proving copyright, especially if property rights to a work have been transferred many times from one owner to another, is difficult and time consuming. All copyright transfer agreements must be submitted to the court. The court must carefully examine each contract individually, in particular the specific scope of the rights transferred under the contract and determine whether the exclusive property rights have actually passed from the author to the current owner. In addition, the transfer of rights is possible to several owners, as well as partially. Given the infinite number of all possible combinations, it becomes increasingly difficult to prove ownership of a work only through human resources or primitive databases. To solve this problem, it will be urgent to develop a special information system that automates all the processes of registration, transfer and proof of copyright, as well as provide a reliable level of protection of the rights to the works of its users.

The **purpose** of the study is to increase the degree of copyright protection and to reduce the cost and timeframe of verifying the authorship or ownership of the work.

To achieve this goal, we have to complete the following **tasks**:

- explore ways to store limited work data;
- explore methods for identifying similar content documents;
- develop a method of outputting digital prints of works that will allow to identify duplicates and near-duplicates;
- develop an information system for registration and confirmation of copyright.

The **object** of study is the state copyright protection process in Ukraine.

The **subject** of the research is the information system of registration and confirmation of copyright.

Scientific novelty of the obtained results. We have developed a method of outputting digital prints of works that allow to detect not only duplicates but also works

close in content by combining the MinHash and Keccak algorithms. Prints of this type are of limited length and will require a limited amount of memory to store in the information system.

Publications. The results of the research were published in the theses of the fourteenth international scientific-practical conference “Mathematical and imitation modeling of systems. MODS '2019', in the abstracts of the Third All-Ukrainian Scientific and Practical Conference of Young Scientists and Students 'Information Systems and Management Technologies' (ISTU-2019).

SYSTEM, COPYRIGHT, HASH FUNCTION, HASHING ALGORITHM,
SIMILAR DOCUMENTS, DIGITAL IMPRINT

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ	11
ВСТУП	13
1 ПРОЕКТНІ РІШЕННЯ З РОЗРОБКИ СИСТЕМИ РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ	16
1.1 Опис бізнес-процесів	16
1.1.1 Опис процесу діяльності	16
1.1.1.1 Реєстрація авторських прав в Україні	16
1.1.1.2 Передача авторського права в Україні	21
1.1.1.3 Підтвердження (доведення) авторського права в Україні	22
1.1.2 Актори і функції	24
1.1.2.1 Автор, його роль та функції у системі	24
1.1.2.2 Власник авторських прав (майнових прав на твір), його роль та функції у системі	25
1.1.2.3 Реєстратор, його роль та функції у системі	25
1.1.2.4 Суд, адвокат та нотаріус. Їх ролі та функції у системі	26
1.1.3 Структура бізнес-процесів	27
1.1.3.1 Процес реєстрації авторського права	27
1.1.3.2 Процес передання авторських прав (прав власності на твір) частково або повністю від однієї фізичної чи юридичної особи до іншої	28
1.1.3.3 Процес підтвердження авторства чи майнових прав на твір	30
1.2 Опис постановки задачі	30
1.3 Рішення з інформаційного забезпечення	32
1.4 Висновки по розділу	35
2 МЕТОДИ ОТРИМАННЯ ЦИФРОВИХ ВІДБИТКІВ ДОКУМЕНТІВ З МОЖЛИВІСТЮ ВИЯВЛЕННЯ БЛИЗЬКИХ ДУБЛІКАТІВ	36

2.1	Постановка задачі на розробку метода отримання цифрових відбитків документів з можливістю виявлення близьких дублікатів	36
2.2	Сучасні алгоритми криптографічного хешування	36
2.2.1	Алгоритм MD5	39
2.2.2	Хеш-функція HAVAL	42
2.2.3	Алгоритм безпечного хешування (Secure Hash Algorithm, SHA)	43
2.2.4	Алгоритм Кессак (SHA-3)	46
2.3	Сучасні методи виявлення подібних за змістом документів	47
2.3.1	Алгоритм шинглів	48
2.3.2	Пошук схожих документів з MinHash	49
2.3.3	Пошук схожих документів з SimHash	49
2.4	Розробка методу отримання цифрових відбитків документів з можливістю виявлення близьких дублікатів	51
2.5	Висновки по розділу	52
3	ОПИС ПРОГРАМНОГО ТА ТЕХНІЧНОГО ЗАБЕЗПЕЧЕННЯ	53
3.1	Засоби розробки	53
3.1.1	Інтегроване середовище розробки (IDE)	54
3.1.2	Інструмент побудови діаграм	55
3.1.3	Інструмент для розробки дизайну та прототипу інтерфейсу користувача	56
3.1.4	Система контролю версій вихідних кодів	57
3.1.5	Хостинг віддаленого сховища Git	58
3.1.6	Інструмент планування та контролю процесу розробки	59
3.2	Архітектура програмного забезпечення	60
3.3	Інструкція користувача	61
3.3.1	Реєстрація твору та авторського права на нього	62
3.3.2	Реєстрація передачі авторського права на твір	63
3.3.3	Підтвердження авторського права на твір	65

3.4	Опис технічного забезпечення	65
3.5	Висновки по розділу	66
4	РОЗРОБКА СТАРТАП-ПРОЕКТУ	68
4.1	Опис ідеї проекту	68
4.1.1	Зміст ідеї, напрямки та вигоди застосування	68
4.1.2	Визначення сильних, слабких та нейтральних характеристик ідеї проекту	69
4.2	Технологічний аудит ідеї проекту	71
4.3	Аналіз ринкових можливостей запуску стартап-проекту	72
4.3.1	Аналіз попиту	73
4.3.2	Потенційні групи клієнтів	73
4.3.3	Аналіз ринкового середовища	74
4.3.4	Аналіз пропозиції	76
4.3.5	Детальний аналіз умов конкуренції в галузі	77
4.3.6	Фактори конкурентоспроможності	78
4.3.7	Аналіз сильних та слабких сторін стартап-проекту	79
4.3.8	SWOT - аналіз стартап - проекту	80
4.3.9	Альтернативи ринкової поведінки	82
4.4	Розроблення маркетингової програми стартап-проекту	82
4.4.1	Визначення стратегії охоплення ринку	82
4.4.2	Базова стратегія розвитку	83
4.4.3	Вибір стратегії конкурентної поведінки	84
4.4.4	Визначення стратегії позиціонування	85
4.5	Розроблення маркетингової програми стартап-проекту	86
4.5.1	Формування маркетингової концепції товару	86
4.5.2	Трирівнева маркетингова модель послуг	87
4.5.3	Визначення меж встановлення ціни	88
4.5.4	Формування системи збуту	89

	10
4.5.5 Концепція маркетингових комунікацій	90
4.6 Висновки по розділу	91
ВИСНОВКИ	92
ПЕРЕЛІК ПОСИЛАНЬ	94
ДОДАТОК А. Графічний матеріал	98
Діаграми способів використання системи для автора, власника майнових прав та інших акторів	99
Діаграма процесу первинної реєстрації твору автором	100
Діаграма процесу передачі майнових прав на твір	101
Діаграма процесу підтвердження авторських та майнових прав на твір	102
Діаграма інформаційної моделі системи реєстрації та підтвердження авторських прав	103
Діаграма класів інформаційної системи реєстрації та підтвердження авторських прав	104
Діаграма розгортання інформаційної системи реєстрації та підтвердження авторських прав	105

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

NIST (National Institute of Standards and Technology) - Національний інститут стандартів і технології, національний орган зі стандартизації у США.

SaaS (Software as a service) - модель поширення програм споживачам, при якій постачальник розробляє веб-програму, розміщує її й керує нею (самостійно або через третіх осіб) з метою використання її замовниками через інтернет.

Блочний шифр - різновид симетричного шифру. Особливістю блочного шифру є обробка блоку декількох байт за одну ітерацію (як правило 8 або 16).

Депонування - процес організованого зберігання чого-небудь.

Криптографічна хеш-функція - це хеш-функція, яка приймає довільний блок даних і повертає рядок встановленого розміру, криптографічне хеш-значення, таке що випадкові або навмисні зміни даних з дуже високою ймовірністю змінять хеш-значення.

Куки (Cookies) - у комп'ютерній термінології поняття, яке використовується для опису інформації у вигляді текстових або бінарних даних, отриманих від веб-сайту на веб-сервері, яка зберігається у клієнта, тобто браузера, а потім відправлена на той самий сайт, якщо його буде повторно відвідано.

Піратство - дії, спрямовані на протиправне використання об'єктів права інтелектуальної власності, що належать іншим особам, умисно вчинені особою, яка розуміє протизаконний характер цих дій, з метою отримання матеріальної вигоди.

Плагіат - привласнення авторства на чужий твір або на чуже відкриття, винахід чи раціоналізаторську пропозицію, а також використання у своїх працях чужого твору без посилання на автора.

Скрін-шот - зображення, отримане комп'ютером, що зображує дійсно те, що бачить користувач на екрані монітора.

Універсальний криптографічний примітив - це чітко визначений криптографічний алгоритм низького рівня, який часто використовуються для

побудови криптографічних протоколів комп'ютерних систем безпеки. Стандартний примітив зазвичай включає в себе стійкі хеш-функції та функції шифрування.

Цифровий відбиток - результат застосування хеш-функції до файлу.

ВСТУП

У зв'язку із швидкими темпами розвитку сфери високих технологій, зростаючою роллю інтелектуальних ресурсів для економіки розвинених країн, на сьогоднішній день значення права інтелектуальної власності і, зокрема, авторського права, важко переоцінити [1].

Варто відзначити, що нині в Україні проблема захисту авторських прав залишається малодослідженою. Зокрема, ані на законодавчому рівні, ані у роз'ясненнях компетентних органів, ані у працях науковців у цій сфері не відображено особливостей належного доказування авторства, наявності авторських прав, вини при розгляді судом справи про порушення авторського права [1].

У зв'язку з цим, доцільним вбачається звернути увагу не тільки на проблеми законодавчого регулювання та державної політики у цій сфері, а й на питання, що виникають в юридичній практиці, а саме під час судового розгляду справ про порушення авторського права [1].

Для того, щоб особі довести свої авторські права, факт порушення таких прав, вину порушника і обґрунтувати розмір компенсації за порушення, потрібно юридично грамотно і своєчасно підійти до збору доказів вказаних обставин, оскільки кожна сторона повинна довести ті обставини, на які вона посилається як на підставу своїх вимог або заперечень, як то передбачено статтею 10 Цивільного процесуального кодексу України (далі – ЦПК України) та статтею 33 Господарського процесуального кодексу України (далі — ГПК України) [1].

В першу чергу, варто наголосити на необхідності доказування наявності у особи повноважень на звернення до суду, оскільки відповідно до статті 1 ГПК України право на звернення з позовом до господарського суду пов'язується з порушенням або оспоруванням прав і охоронюваних законом інтересів позивача. Те саме зазначено у статті 3 ЦПК України [1].

Крім того, положенням частини 3 пункту 12 Постанови Пленуму Верховного Суду України (далі – ВСУ) від 04.06.2010 р. №5 “Про застосування судами норм законодавства у справах про захист авторського права і суміжних прав”, передбачено, що позивач повинен довести факт наявності в нього авторського права [1].

Враховуючи вказану норму, суду спочатку необхідно з’ясувати обсяг авторських майнових прав особи, яка подає позовну заяву, оскільки від цього залежить її процесуальна правоздатність [1].

Слід зазначити, що це, зокрема, стосується випадків, коли позов подається представником автора (посередником) – певною організацією, яка уклала договір із автором, або організацією, яка уклала договір з іншою організацією, яка має договір з автором [1].

Зокрема, на практиці мають місце судові справи про порушення авторських прав за позовом організацій, які набули авторських прав через укладення договору із такою ж організацією, яка є посередником автора [1].

В такому разі позивач має надати суду весь ланцюг договорів від автора до позивача і встановити конкретний обсяг прав, що передається за цим договором. Слід зазначити, що така позиція висловлена у постанові Київського апеляційного господарського суду від 17 липня 2013 року у справі № 910/4773/13, якою скасовано рішення суду першої інстанції та відмовлено у задоволенні позову на тій підставі, що у справі були відсутні належні правові докази наявності у позивача майно. Більше того, в даній справі суд першої інстанції дослідив весь ланцюг договорів і лише після цього було винесено рішення [1].

Аналогічної думки притримується і Господарський суд м. Києва у постанові від 22.10.2008р. у справі №20/194 за позовом ПП “Українське музичне видавництво” до представництва “Фудеволд маркетинг ППТІ ЛТД” про стягнення 135 000 грн. за порушення виключних майнових авторських прав. Даним рішенням суд, дослідивши увесь ланцюг договорів від автора до позивача, відмовив у задоволенні позовних

вимог, на тій підставі, що у Позивача були відсутні належні авторські права. Слід зазначити, що дане рішення залишене без змін Київським апеляційним господарським судом та Вищим господарським судом України [1].

Варто додати, що окрім того, що в матеріалах справи мають міститися усі договори про передання авторських прав, суд має ретельно дослідити кожний договір окремо, зокрема, конкретний обсяг прав, що передається за договором, та встановити, чи дійсно перейшли виключні майнові права від автора до нинішнього Позивача [1].

Як бачимо, для спрощення процесів реєстрації договорів про передання авторських прав та доведення прав за допомогою аналізу ланцюжка їх передач від автора до поточного власника необхідно створити автоматизовану систему, яка візьме на себе усі складні функції, пов'язані з описаними вище процесами.

Отже, потрібна система, яка надаватиме можливості первинної реєстрації творів авторами, реєстрації передачі авторських прав між фізичними та юридичними особами та підтвердження (доведення) авторських прав чи прав власності на твори.

1 ПРОЕКТНІ РІШЕННЯ З РОЗРОБКИ СИСТЕМИ РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ

1.1 Опис бізнес-процесів

1.1.1 Опис процесу діяльності

Система захисту авторських прав в Україні складається з наступних процесів: реєстрація авторських прав, передання їх частково чи повністю від однієї фізичної чи юридичної особи до іншої, доведення авторства чи права власності на твір.

1.1.1.1 Реєстрація авторських прав в Україні

Авторське право, як і будь-який інший об'єкт права інтелектуальної власності, має охоронятися як на світовому рівні, так і на рівні конкретної країни. Одним із найважливіших важелів охорони авторського права, що надається державою, є його реєстрація.

Реєстрація авторських прав в Україні — це процес отримання охоронного документа, а саме “свідоцтва про реєстрацію авторських прав на твір” шляхом державної реєстрації [2].

Під твором мається на увазі будь-яке матеріальне втілення оригінальної форми вираження Автора. Іншими словами, в якості твору можуть виступати наступні об'єкти авторських прав: музика, пісні, вірші, комп'ютерні програми, переклади, фонограми і т.д [2].

Сьогодні в Україні найбільш затребувана реєстрація авторських прав на комп'ютерні програми, дизайн сайту, контент сайту, пісні, музику, вірші, книги [2]. Комп'ютерні програми охороняються Законом як літературні твори. Крім реєстрації авторських прав, користується популярністю і реєстрація передачі авторських прав на перераховані вище об'єкти авторського права [2].

Реєстрація авторських прав в Україні регулюється Цивільним кодексом і Законом України «Про авторське право і суміжні права», а також Бернською

конвенцією про охорону літературних і художніх творів (п. 2 ст. 5). Але основним законодавчим актом України щодо авторського права та суміжних прав є Закон України "Про авторське право і суміжні права" в редакції від 22.05.2003 року [3].

Цей Закон визначає наступний перелік об'єктів авторського права:

- літературні письмові твори белетристичного, публіцистичного, наукового, технічного або іншого характеру (книги, брошури, статті тощо);
- виступи, лекції, промови, проповіді та інші усні твори;
- комп'ютерні програми;
- бази даних;
- музичні твори з текстом і без тексту;
- драматичні, музично-драматичні твори, пантоміми, хореографічні та інші твори, створені для сценічного показу, та їх постановки;
- аудіовізуальні твори;
- твори образотворчого мистецтва;
- твори архітектури, містобудування і садово-паркового мистецтва;
- фотографічні твори, у тому числі твори, виконані способами, подібними до фотографії;
- твори ужиткового мистецтва;
- ілюстрації, карти, плани, креслення, ескізи, пластичні твори, що стосуються географії, геології, топографії, техніки, архітектури та інших сфер діяльності;
- сценічні обробки творів, обробки фольклору, придатні для сценічного показу;
- похідні твори;
- збірники творів, збірники обробок фольклору, енциклопедії та антології, збірники звичайних даних, інші складені твори за умови, що вони є результатом творчої праці;
- тексти перекладів;

- інші твори [3].

Згідно із законодавством в Україні охороняються як оприлюднені твори, так і не оприлюднені, завершені та незавершені. Авторське право на будь-який твір виникає внаслідок факту його створення та вираження його в матеріальному або нематеріальному носії. Ці положення відображені у національному законодавстві, зокрема, у пункті другому статті 11 Закону України “Про авторське право і суміжні права”. Для виникнення і здійснення авторського права не вимагається реєстрація твору. Отже, кожний автор, створивши твір, автоматично набуває прав власності на нього [3].

При тому, що реєструвати авторські права на твір не обов’язково, авторів у багатьох випадках важливо подбати про фіксацію дати створення свого твору. Адже недооцінка цього моменту в майбутньому може ускладнити авторів захист своїх прав на власний твір. Тому одним із засобів фіксації дати створення твору є його реєстрація та депонування — процес організованого зберігання [1].

В Україні, як і в інших державах передбачений механізм реєстрації прав автора. Реєстрація авторського права дає авторів твору наступні переваги:

- в спірних моментах звільняє автора від необхідності доводити своє авторство, так як з отриманням свідоцтва у нього автоматично виникає презумпція; отримана презумпція може служити звільненням від доказу авторства на твір не тільки в Україні, але також в інших країнах світу;
- закріплює і підтверджує для третіх осіб володіння майновими правами конкретного автора;
- дозволяє укладати ліцензійні договори про передачу авторського права;
- дозволяє продавати право власності учасника його авторського права на створений твір;
- дозволяє вносити автору створений твір в статутний капітал підприємства [2].

Заявки на реєстрацію авторського права подаються в Українське агентство з авторських і суміжних прав. Для реєстрації авторських прав на твір необхідно подати заявку, яка складається з наступних документів:

- заява, викладена українською мовою, що складається за формою затвердженою Державною службою інтелектуальної власності;
- примірник твору (у матеріальній формі, визначеній пунктами 3, 15 Порядку);
- документ, що свідчить про факт і дату оприлюднення твору (за наявності);
- документ або копія документа про сплату збору за підготовку до реєстрації авторського права, або копія документа, що підтверджує наявність пільг;
- документ про сплату збору за оформлення та видачу свідоцтва або копія документа, що підтверджує наявність пільг [4].

До заяви на реєстрацію авторського права на комп'ютерну програму, крім зазначених документів, додається настанова щодо використання програми [2].

До заяви на реєстрацію авторського права на базу даних, крім зазначених документів, додається настанова щодо використання бази даних і опис структури бази даних [2].

До заяви на реєстрацію авторського права на твори архітектури, містобудування, садово-паркового та монументального мистецтва, крім зазначених документів, додається анотація, яка повинна містити назву твору, його місцезнаходження (адресу), архітектурні характеристики параметри, час і місце створення [2].

Крім цього, повинен бути сплачений державний збір за подачу заявки в установленому розмірі [2].

Заява підписується автором або особою, яка має авторське право. Подати заяву на реєстрацію авторського права може автор, його спадкоємець або їх правонаступник, як фізична, так і юридична особа, як самотійно, так і через

довірену особу. В останньому випадку до матеріалів заявки додається довіреність, видана заявником на ведення діловодства [2].

У разі, якщо авторське право оформляється на особу, яка не є автором (на твір, створений в результаті службового завдання), то необхідно надати документ, що підтверджує приналежність майнових прав цієї особи (ст. 429 Цивільного кодексу України) [2].

Розгляд заявки на реєстрацію авторського права передбачає тільки перевірку наявності та правильності оформлення всіх необхідних документів. Після отримання позитивного рішення за заявкою заявник повинен оплатити державне мито за видачу свідоцтва в установленому розмірі [2].

Після оплати відомості про видачу свідоцтва на авторське право публікуються в офіційному бюлетені Департаменту інтелектуальної власності, заносяться до відповідного Державного реєстру, а заявник одержує свідоцтво про реєстрацію авторського права на твір [2].

Реєстрація авторського права займає приблизно 2-3 місяці. Держмитом на теперішній час становить:

- для фізичних осіб — 63,75 грн;
- для юридичних осіб — 187 грн [2].

За умови, що всі документи оформлені правильно, термін розгляду починає відраховуватися з дня надходження заявки і становить один календарний місяць [2].

Негативне рішення може бути прийняте через неправильно заповнені документи і заявку, неповний пакет документів, неналежно оформлений зразок або неможливість патентування в зв'язку з невідповідністю об'єкта вимогам законодавства. В цьому випадку у відомство направляється відповідь з аргументованим запереченням, якщо справа у оформленні — помилки виправляються, і готується переподача [2].

1.1.1.2 Передача авторського права в Україні

Відповідно до ч.ч. 3,4 ст. 32 Закону України “Про авторське право та суміжні права” за авторським договором про передачу виключного права на використання твору автор (чи інша особа, яка має виключне авторське право) передає право використовувати твір певним способом і у встановлених межах тільки одній особі, якій ці права передаються, і надає цій особі право дозволяти або забороняти подібне використання твору іншим особам [1].

За авторським договором про передачу невиключного права на використання твору автор (чи інша особа, яка має авторське право) передає іншій особі право використовувати твір певним способом і у встановлених межах [1].

Отже, виходячи із змісту вказаних вище положень, право дозволяти або забороняти подібне використання твору іншим особам, і, як наслідок, звертатись до суду із позовом про захист цього права і стягнення компенсації за порушення такого права, належить лише тій особі, яка має виключне право на використання твору [1].

Тобто, якщо у договорі, на який посилається позивач як на доказ, що підтверджує його повноваження, зазначено про те, що йому передано невиключне право на використання твору, в такому разі звертатися до суду він не має права, оскільки його права не є порушеними [1].

Також, варто наголосити на тому, що єдиним належним доказом наявності у позивача виключного авторського права є авторський договір, а не будь-який інший документ (довідка, лист, повідомлення і т.д.), навіть якщо він засвідчений нотаріально або іншим чином. Це пояснюється наступним [4].

Відповідно до ч.ч. 1-3 статті 32 Закону України “Про авторське право та суміжні права” автору та іншій особі, яка має авторське право, належить виключне право надавати іншим особам дозвіл на використання твору будь-яким одним або всіма відомими способами на підставі авторського договору [4].

Використання твору будь-якою особою допускається виключно на основі авторського договору, за винятком випадків, передбачених ст.ст. 21-25 Закону.

Передача права на використання твору іншим особам може здійснюватися на основі авторського договору про передачу виключного права на використання твору або на основі авторського договору про передачу невиключного права на використання твору [4].

Статтею 1107 ЦК України передбачено, що розпорядження майновими правами інтелектуальної власності здійснюється на підставі таких договорів:

- ліцензія на використання об'єкта права інтелектуальної власності;
- ліцензійний договір;
- договір про створення за замовленням і використання об'єкта права інтелектуальної власності;
- договір про передання виключних майнових прав інтелектуальної власності;
- інший договір щодо розпорядження майновими правами інтелектуальної власності [1].

Таким чином, маємо констатувати, що на підтвердження своїх повноважень позивач має надати саме договір про передання авторських прав, а не будь-який інший документ, оскільки трапляються випадки, коли позивачі як на підтвердження своїх повноважень посилаються на апостильовані довідки, видані організаціями, що уклали авторський договір із автором і про що зазначається у такій довідці. Варто зазначити, що в даному випадку належним доказом підтвердження повноважень позивача має бути не довідка від організації, із якою позивач уклав договір, а сам авторський договір організації, на яку посилається позивач, із автором твору [1].

1.1.1.3 Підтвердження (доведення) авторського права в Україні

Хоча зареєстровані та незареєстровані твори, згідно Закону охороняються однаково, реєстрація авторського права на твір дає можливість отримати документ - свідоцтво, який підтверджує належність авторських прав тій чи іншій особі. Свідоцтво є єдиним переконливим засобом підтвердження свого авторства на твір у

разі його неправомірного використання іншими особами, в судових справах, при укладанні угод на отримання винагороди за використання твору тощо [3].

Існують різні види порушень авторських прав. До них відноситься вчинення будь-яких дій, що порушують особисті немайнові та майнові права автора. Всім відомі піратство й плагіат. Особливо це стосується різних публікацій на інтернет-сторінках чи у соціальних мережах [5].

Законом передбачена цивільна та навіть кримінальна відповідальність у разі порушення авторських прав. Дійсно, до суду автору потрібно надати докази того, що твір належить саме йому, зібрати докази порушення та обґрунтувати розмір шкоди, якої він зазнав [5].

На сьогодні не існує однозначної відповіді стосовно того, чи прийме суд скрін-шот інтернет-сторінки як доказ авторства. Можна прохати суд оглянути веб-сторінку прямо під час судового засідання. Таким чином, суд зафіксує, що саме автор має необхідні ключі для входу до адміністративної частини сайту і має змогу робити публікації та редагувати їх. Якщо цього буде недостатньо, варто звернутись до адміністрації соціальної мережі із запитом IP-адреси, з якої опубліковано статтю. Також можна замовити комп'ютерно-технічну експертизу [5].

Варто зазначити, що на сьогодні не існує конкретного судового прикладу доведення авторства публікації на веб-сторінці. Процес дуже складний та непередбачуваний [5].

В Україні не надто розповсюджена практика судових позовів задля захисту власного авторського права. Це пов'язано і з недовірою до судової системи і страхом. Крім того, українське законодавство у цій галузі має багато недоліків, тому найбільшою проблемою залишається підтвердження права на інтелектуальну власність. У суді також потрібно довести, власне, факт порушення. Це становить велику проблему, адже в Україні немає конкретного алгоритму дій для таких випадків.

1.1.2 Актори і функції

В описаних процесах реєстрації, передачі та підтвердження авторських прав можна виділити такі ролі (таких акторів): автор, власник авторських прав (майнових прав на твір), реєстратор, суд, адвокат, нотаріус. Цей перелік може бути доповнено, але на даному етапі обмежимося ним.

1.1.2.1 Автор, його роль та функції у системі

Автор створює будь-який твір та подає заявку разом з примірником твору на реєстрацію авторських прав на нього. Автор може одночасно виконувати і роль власника майнових прав на твір. Ще однією функцією Автора в системі являється запит на підтвердження (доведення) його авторства та майнових прав на певний твір. Описані функції автора зображено на наступній діаграмі (рисунок 1.1).

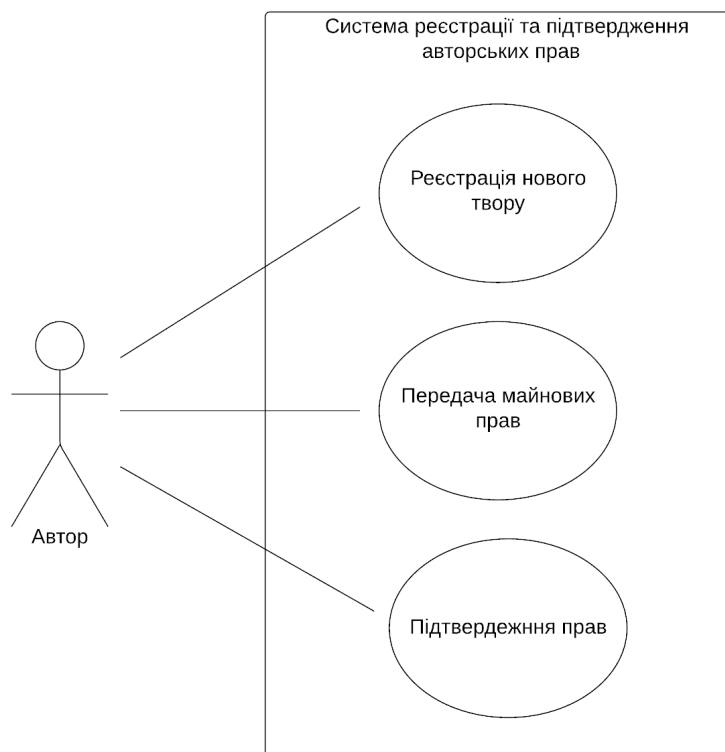


Рисунок 1.1 - Діаграма способів використання системи для автора

1.1.2.2 Власник авторських прав (майнових прав на твір), його роль та функції у системі

Власник авторських прав (майнових прав на твір) реєструє договір про повну або часткову передачу цих прав іншій фізичній чи юридичній особі, внаслідок чого право власності на твір (авторське право) переходить до другої сторони договору.

Власник авторських прав може одночасно виконувати і роль Автора, який створив об'єкт інтелектуальної власності.

Власник авторських прав (майнових прав на твір) може подати заявку на підтвердження (доведення) його майнових прав на певний твір за допомогою системи.

Описані способи використання системи власником майнових прав зображено на наступній діаграмі (рисунок 1.2).

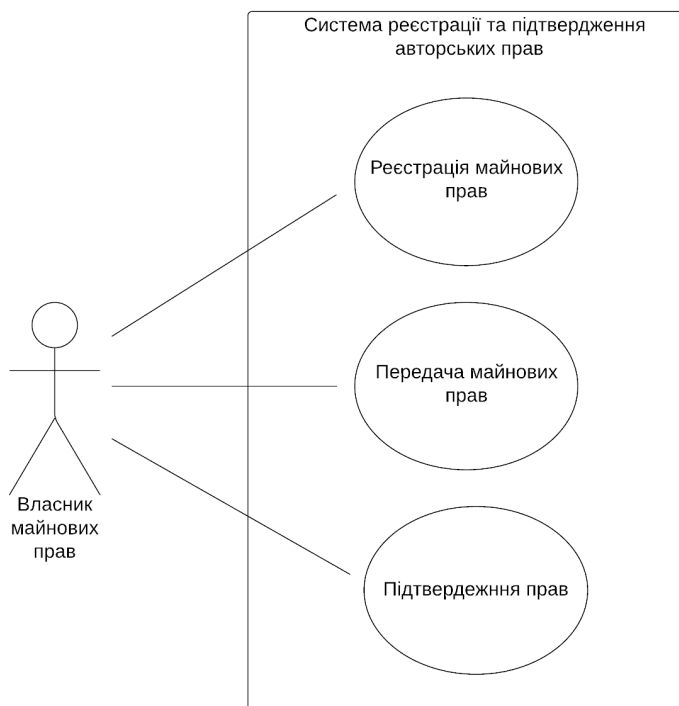


Рисунок 1.2 - Діаграма способів використання системи для власника майнових прав

1.1.2.3 Реєстратор, його роль та функції у системі

Реєстратор приймає заявку та пакет документів на реєстрацію авторських прав Автором, перевіряє правильність оформлення документів та належність твору особі,

що подає заявку на реєстрацію, та реєструє твір і авторське право заявника у системі.

Реєстратор приймає заявку на реєстрацію договору про часткову або повну передачу прав на інтелектуальну власність від власників таких прав, що бажають передати його іншій стороні внаслідок купівлі-продажу або інших умов, перевіряє дійсність договору та реєструє його в системі разом з передачею прав власності на твір за іншою стороною.

Реєстратор здатний обробити запит від Автора, Власника авторського права, Суду, Адвоката або Нотаріуса щодо підтвердження авторського права (майнового права) на твір за допомогою перевірки реєстраційних даних у базі даних системи.

Функції реєстратора можуть бути покладені на інформаційну систему, якщо автоматизувати усі дії та операції, що описані вище.

1.1.2.4 Суд, адвокат та нотаріус. Їх ролі та функції у системі

Суд, адвокат та нотаріус виконують кожен свою роль, але однакову функцію у процесі доведення авторського права.

Суд приймає рішення щодо визнання авторського права за певною особою, яка заявляє, що являється автором або власником прав на твір. Для цього Суд звертається із запитом про підтвердження цього факту до системи через Реєстратора.

Адвокат допомагає автору або власнику твору довести своє право перед судом. Він здійснює це через звернення із запитом про підтвердження цього факту до системи через Реєстратора.

Нотаріус допомагає автору або власнику твору перевірити дійсність авторства чи права власності над твором певної особи, що звертається до нього. Він підтверджує реєстрацію права, виконуючи запит до системи через Реєстратора.

Описані способи використання системи такими авторами як суд, адвокат, нотаріус зображено на наступній діаграмі (рисунок 1.3).

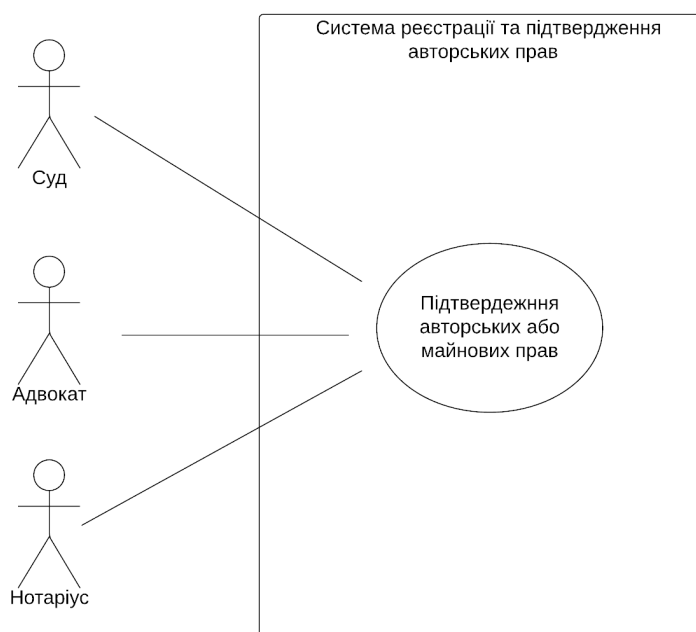


Рисунок 1.3 - Діаграма способів використання системи судом, адвокатом, нотаріусом

1.1.3 Структура бізнес-процесів

Наведемо послідовність дій в межах процесів системи реєстрації та підтвердження авторських прав, а саме: реєстрації авторських прав, передання їх частково або повністю від однієї фізичної чи юридичної особи до іншої, підтвердження авторства чи права власності на твір.

1.1.3.1 Процес реєстрації авторського права

Процес реєстрації авторського права починається з моменту подання автором заявки на реєстрацію разом з примірником твору та іншими необхідними документами.

Після подачі пакету документів вони перевіряються на правильність заповнення та дійсність вказаних даних. Наприклад, можуть перевірятись дійсність вказаних у документах персональних та реєстраційних даних заявника (автора),

дійсність оприлюднення твору (якщо це мало місце до подання заявки), здійснення оплати зборів (мита) чи наявності пільг, існування у системі раніше зареєстрованого твору, ідентичного поданому на реєстрацію.

Процес завершується успішною реєстрацією твору та його автора разом з фіксацією мітки часу з подальшою видачею свідоцтва про реєстрацію, або відмовою у реєстрації внаслідок виникнення невідповідності вимогам під час перевірки твору та документів на попередньому кроці процесу (рисунок 1.4).

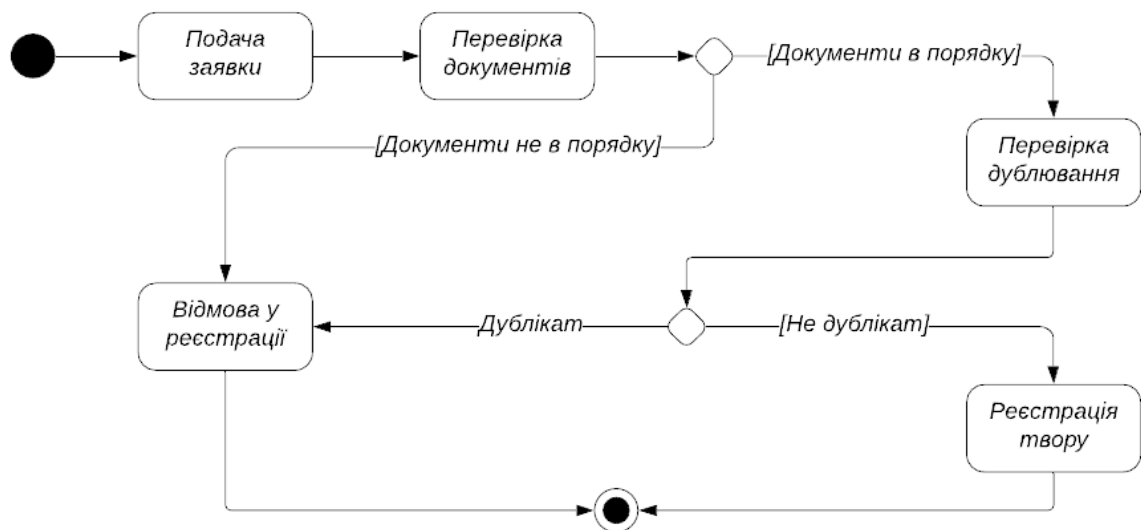


Рисунок 1.4 - Діаграма процесу первинної реєстрації твору автором

Даний процес може бути покращений за рахунок автоматизації прийому та перевірки документів, підтвердження, що в базі даних системи не зареєстровано ідентичного твору. Крім того, перевірка унікальності твору може бути покращена за рахунок перевірки не тільки наявності ідентичного твору, але й дуже подібного (частково зміненого). Для цього необхідно розробити новий спосіб реєстрації творів і, як наслідок, нову автоматизовану систему.

1.1.3.2 Процес передавання авторських прав (прав власності на твір) частково або повністю від однієї фізичної чи юридичної особи до іншої

Процес передавання авторських прав (прав власності на твір) частково або повністю від однієї фізичної чи юридичної особи до іншої починається від моменту подання заявки на реєстрацію договору про передачу (відчуження) майнових прав на твір або передачу виключного (невиключного) права на використання твору.

Після подачі пакету документів вони перевіряються на правильність заповнення та дійсність вказаних даних. Наприклад, можуть перевірятись дійсність вказаних у документах персональних та реєстраційних даних, існування у системі твору, права на який передається згідно договору, та належність твору особі, що передає права на нього.

Процес завершується успішною реєстрацією передачі прав на твір та його нового власника разом з фіксацією мітки часу з подальшою видачею свідоцтва про реєстрацію, або відмовою у реєстрації внаслідок виникнення невідповідності вимогам під час перевірки твору та документів на попередньому кроці процесу (рисунок 1.5).

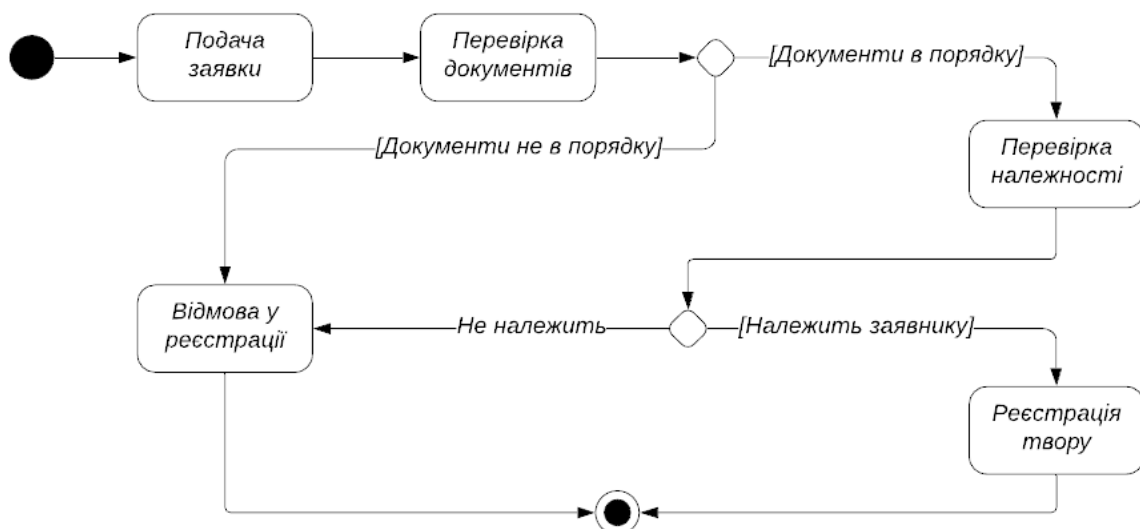


Рисунок 1.5 - Діаграма процесу передачі майнових прав на твір

Даний процес може бути покращений за рахунок автоматизації прийому та перевірки документів, підтвердження того, що в базі даних системи раніше було зареєстровано твір та що він належить стороні, яка передає права на нього.

1.1.3.3 Процес підтвердження авторства чи майнових прав на твір

Процес підтвердження авторства чи майнових прав на твір починається з моменту виконання автором чи власником твору запиту на підтвердження авторства чи майнових прав на твір.

Після приймання запиту відбувається пошук відповідної реєстраційної інформації у базі даних системи, що доводить авторство чи право власності на твір.

Процес завершується разом з наданням інформації про реєстрацію авторських прав на твір, або реєстрацію виключного (невиключного) права власності на твір разом з усім ланцюжком передач за договорами, що підтверджує передання прав від автора, що створив об'єкт інтелектуальної власності, до кінцевого власника (рисунок 1.6).

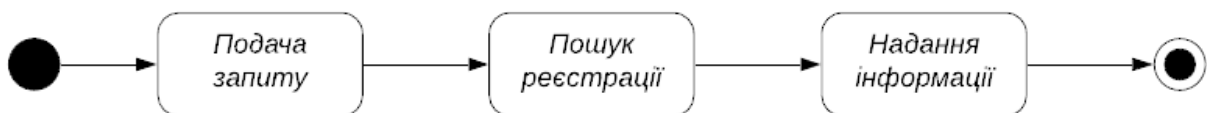


Рисунок 1.6 - Діаграма процесу підтвердження авторських та майнових прав на твір

Описаний вище процес являється вже удосконаленим, оскільки в реальному стані системи захисту авторських прав в Україні дуже складно за один запит отримати всю необхідну інформацію та дослідити увесь ланцюг передачі авторських прав на твір до кінцевого власника.

1.2 Опис постановки задачі

Проаналізувавши існуючі бізнес-процеси української системи захисту авторських прав та майнових прав на твори, зокрема реєстрацію авторства, передачу майнових прав на інтелектуальну власність та можливості підтвердження прав за допомогою неї, рівень автоматизації зазначених процесів та можливості щодо їх покращення, приходимо до висновку, що процеси можна і потрібно покращити та автоматизувати за допомогою нової інформаційної системи.

Нова система надаватиме можливості первинної реєстрації творів авторами, реєстрації передачі авторських прав між фізичними та юридичними особами та підтвердження (доведення) авторських прав чи прав власності на твори.

Метою створення нової системи являється підвищення рівня захисту авторських прав фізичних та юридичних осіб. Головними вимогами до нової системи є зменшення вартості та часових рамок підтвердження своїх авторських прав чи прав власності на твір, а також можливість швидко та легко віднайти весь ланцюг договорів про передання прав від автора до позивача і встановити конкретний обсяг прав, що передається за цими договорами [14].

Система повинна забезпечити збереження цифрового відбитку твору замість збереження оригінального контенту файлу. Це дасть змогу використовувати менший розмір пам'яті в файлових сховищах. Крім того система повинна надавати змогу виявляти дублікати (вже раніше зареєстровані твори, що є абсолютною копією), та дуже близькі до оригінального зареєстрованого твору. Копію знайти дуже просто, оскільки однаковий зміст дасть один і той самий цифровий відбиток, якщо використовувати для цього криптографічні хеш-функції. Але якщо змінити лише один символ у творі, то за допомогою звичайних хеш-функцій ми отримаємо вже зовсім інший відбиток твору. Отже, потрібно розробити такий метод хешування, при якому по цифровому відбитку твору, поданому на реєстрацію, можна знайти в системі відбиток такого самого чи дуже близького за змістом твору.

Для досягнення мети потрібно виконати наступні завдання:

- а) проаналізувати сучасні алгоритми криптографічного хешування;
- б) дослідити методи виявлення подібних за змістом документів;
- в) розробити метод виведення цифрових відбитків творів, які будуть мати однаковий фіксований розмір та дадуть змогу виявляти не тільки дублікати, але й близькі за змістом твори;
- г) розробити інформаційну систему реєстрації та підтвердження прав, складовими якої є:
 - 1) реєстрація творів і авторства на них;
 - 2) реєстрацію передавання майнових прав на твори;
 - 3) підтвердження авторства шляхом видачі ланцюга передачі прав від моменту реєстрації до поточного власника.

1.3 Рішення з інформаційного забезпечення

Виходячи з вимог до інформаційної системи реєстрації та підтвердження авторських прав, вона повинна надати користувачу змогу отримати інформацію про поточну належність твору певному власнику, а також весь ланцюг передачі часткових та повних майнових прав від моменту реєстрації твору.

Вхідними даними системи для реєстрації твору являються:

- прізвище, ім'я та по батькові автора або назва юридичної особи;
- ідентифікаційний податковий номер або реєстраційний номер за ЄДРПОУ;
- дата створення твору;
- файл-копія твору;
- дата та місце оприлюднення твору.

Вхідними даними системи для передавання майнових прав на твір являються:

- файл-копія твору або його хеш;
- дані фізичної (ім'я та індивідуальний податковий номер) або юридичної (назва та реєстраційний номер) особи-приймача прав.

Вхідними даними системи для підтвердження авторства чи майнових прав на твір являється файл-копія твору або його хеш.

Вихідними даними системи для реєстрації нового твору та передавання майнових прав нього буде сертифікат - електронне свідоцтво того, що операцію здійснено успішно та особа чи організація, указана на ньому, являється поточним власником твору.

Вихідними даними для операції підтвердження власності або авторства являється весь ланцюг передачі прав, починаючи від останньої операції, закінчуючи первинною реєстрацією. У разі наявності передач прав декільком власникам, структура вихідних даних вже буде деревоподібною. Якщо операція була лише одна - реєстрація твору - то вона і буде повернута із системи.

Основною сутністю інформаційної моделі визначимо транзакцію - реєстрацію твору або передачу майнових прав на нього від одного власника до іншого. Транзакція містить посилання на попереднього та поточного власників у вигляді їх ідентифікаторів після виконання передачі прав або реєстрації твору та на сутність, що описує твір.

Додатково створимо сутність, що дозволяє створювати та зберігати сертифікати - електронні документи, що доводять належність творів певним власникам.

Усі описані відношення між основними сутностями зображені на діаграмі зв'язків між основними сутностями нашої інформаційної системи (рисунок 1.7).

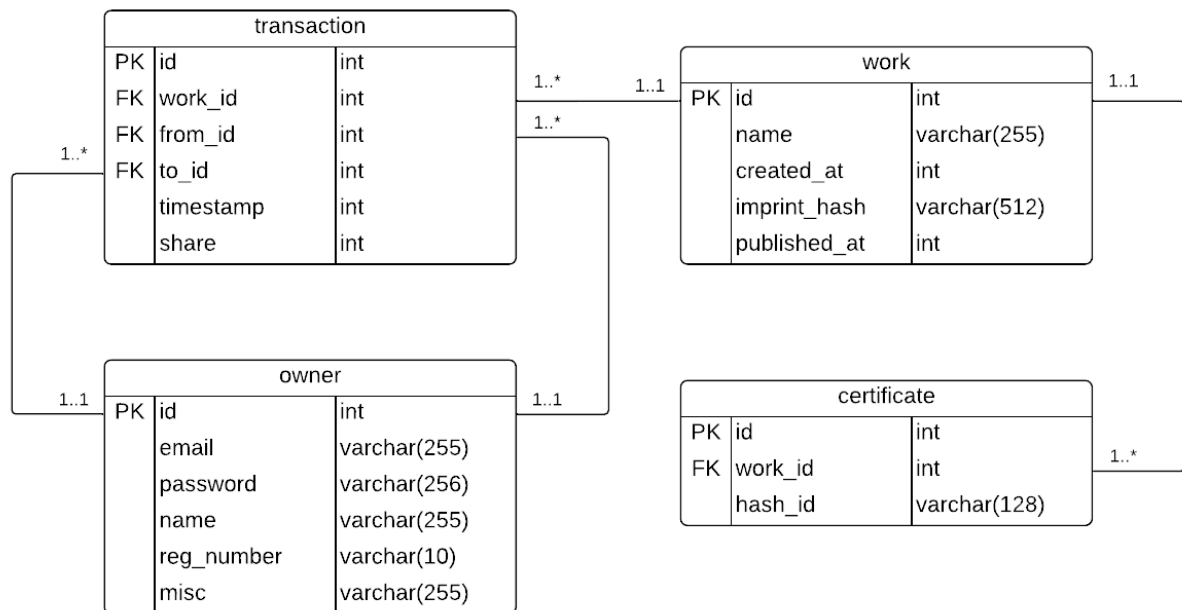


Рисунок 1.7 - Діаграма інформаційної моделі системи реєстрації та підтвердження авторських прав

На основі визначених сутностей та діаграми інформаційної моделі системи (див. рис. 1.7) опишемо основні таблиці її реляційної бази даних.

Таблиця транзакцій. Вона буде складатись з наступних стовпчиків:

- унікальний ідентифікатор транзакції (число, основний ключ, на його основі буде створено індекс);
- ідентифікатор твору (число, зовнішній ключ, посилання на таблицю творів);
- ідентифікатор передавача прав (число, зовнішній ключ, посилання на таблицю власників);
- ідентифікатор приймача прав (число, зовнішній ключ, посилання на таблицю власників);
- мітка часу (число, час здійснення транзакції);
- частка переданих майнових прав (число, кількість відсотків).

Таблиця власників. Її рядки будуть містити наступні поля:

- унікальний ідентифікатор (число, основний ключ, на його основі буде створено індекс);

- електронна пошта (символьний рядок, унікальне значення);
- пароль (символьний рядок у вигляді хешу паролю користувача);
- ім'я або назва (символьний рядок);
- реєстраційний номер або ППН (символьний рядок, дозволяються лише цифри);
- додаткова інформація (доручення, довірена особа та ін.).

Таблиця творів. Записи цієї таблиці складатимуться з наступних полів:

- унікальний ідентифікатор (число, претендент на індекс);
- назва твору (символьний рядок);
- дата створення (число, мітка часу);
- хеш твору (символьний рядок);
- дата публікації (число, мітка часу);
- додаткова інформація (місце публікації та ін.).

Таблиця сертифікатів. Складатиметься з наступних стовпчиків:

- унікальний ідентифікатор (число, претендент на індекс);
- ідентифікатор твору (число, посилання на таблицю творів);
- хеш-ідентифікатор для перегляду сертифікату по прямому посиланню на нього (символьний рядок).

1.4 Висновки по розділу

В даному розділі досліджено існуючі процеси реєстрації та підтвердження авторських прав в системі державного захисту авторських прав в Україні. У ході виконання аналізу покращено усі процеси за рахунок автоматизації їх основних частин, а особливо процес підтвердження права на твір з віднайденням всього ланцюга часткової та повної передачі прав від реєстрації до поточного власника (або власників). Розробка нової системи реєстрації та підтвердження авторських прав підвищить рівень їх захисту та імплементує у собі рішення проблем доведення авторських прав.

2 МЕТОДИ ОТРИМАННЯ ЦИФРОВИХ ВІДБИТКІВ ДОКУМЕНТІВ З МОЖЛИВІСТЮ ВИЯВЛЕННЯ БЛИЗЬКИХ ДУБЛІКАТІВ

2.1 Постановка задачі на розробку метода отримання цифрових відбитків документів з можливістю виявлення близьких дублікатів

Потрібно розробити такий метод хешування, при якому можна отримати відбиток (хеш) будь-якого твору, за яким можна буде знайти дублікат або близький за змістом твір, і який буде мати обмежену та фіксовану довжину для оптимального зберігання в базі даних та унеможливить розшифровку інформації, тобто знайдення змісту твору за допомогою розшифровки хешу [15]. Проаналізуємо сучасні алгоритми криптографічного хешування та методи виявлення подібних за змістом документів та спробуємо знайти оптимальне та краще рішення для виконання поставленого завдання.

2.2 Сучасні алгоритми криптографічного хешування

Функція хешування – це відображення, що перетворює повідомлення довільної довжини у хеш-код фіксованої довжини n , тобто $F : \{0, 1\}^* \rightarrow \{0, 1\}^n$. Криптографічна хеш-функція повинна мати додаткові властивості, основні з яких: стійкість до слабкої колізії, стійкість до сильної колізії та стійкість до побудови прообразу. Більшість сучасних криптографічних протоколів використовують функції хешування. Як правило, надійність таких функцій є критичною для надійності усього протоколу. Оскільки сам хеш-код певною мірою ідентифікує деяку інформацію, то хеш функція повинна забезпечувати конфіденційність цієї інформації (бути стійкою до пошуку прообразу) та цілісність інформації (слабка та сильна стійкість до колізії). Варто зазначити, що стійкість до пошуку прообразу є найбільш критичною властивістю криптографічної функції хешування, оскільки всі інші атаки можна легко звести до пошуку прообразу.

Для криптографічних хеш-функцій також важливо, щоб при найменшій зміні аргументу значення функції сильно змінювалося (лавинний ефект). Зокрема, значення хешу не повинно давати витоку інформації навіть для окремих бітів аргументу. Ця вимога є запорукою криптостійкості алгоритмів хешування, які хешують пароль користувача для отримання ключа [29].

Хеш-функції також використовуються як перевірка деякої інформації на ідентичність оригіналу, без використання оригіналу. Для звірки використовується хеш-значення перевіреної інформації [30].

Розглянемо сучасні алгоритми хешування.

До алгоритмів сімейства MD (Message Digest) входять алгоритми MD2, MD4, MD5 та MD6. Перші три з них широко використовувалися і продовжують широко використовуватися, незважаючи на те, що існують відомі криптоаналітичні атаки проти кожного з них [7].

Алгоритми MD мають різні характеристики. Вони розроблені в різні роки відомим криптологом Рональдом Рівестом (Ronald Rivest) за участю деяких інших фахівців RSA Laboratories - наукового підрозділу компанії RSA Data Security Inc.

MD2 – хешує дані змінного розміру з отриманням 128-бітного вихідного значення. Перш за все, виконується доповнення вхідних даних до розміру, кратного 16 байтам. Потім обчислюється 16-байтна контрольна сума повідомлення, яка також доповнює повідомлення перед його подальшою обробкою [31].

MD4 – структура алгоритму повністю відповідає схемі Меркле-Дамгаарда. Перш за все, вхідне повідомлення розбивається на блоки розміром по 512 бітів. Останній блок завжди доповнюється до 512-бітового розміру [32].

MD5 – 128-бітовий алгоритм хешування. Є покращеною в плані безпеки версією MD4 [33].

MD6 – алгоритм хешування змінної розрядності. Пропонується на зміну менш досконалому MD5 [34].

В ході проведених досліджень вченими було виявлено цілу низку прогалин і вразливостей в алгоритмі MD5. Основною причиною слабкої захищеності MD5 є відносно легке знаходження колізій при шифруванні [7].

Під колізією розуміють можливість отримання однакового результату обчислень хеш-функції при різних вхідних значеннях.

Secure Hash Algorithm 1 (SHA-1) – використовується в багатьох криптографічних додатках і протоколах. Також рекомендований в якості основного для державних установ у США. Принципи, покладені в основу SHA-1, аналогічні тим, які використовувалися Рональдом Рівестом при проектуванні MD4 [16].

SHA-2 – збірна назва односпрямованих хеш-функцій SHA-224, SHA-256, SHA-384 і SHA-512. Хеш-функції призначені для створення “відбитків” або “дайджестів” повідомлень довільної бітової довжини. Застосовуються в різних додатках або компонентах, пов'язаних із захистом інформації [17].

ЛН - Сімейство з чотирьох криптографічних функцій: ЛН-224, ЛН-256, ЛН-384 і ЛН-512. Алгоритм побудований за схемою Меркле-Дамгаарда. Вхідне повідомлення розбивається на блоки по 512 біт. Можливі розміри хеш-значень - 224, 256, 384 і 512 біт. Незалежно від розміру вихідного значення алгоритму в ньому використовується одна і та ж функція стиснення, що формує після обробки останнього блоку повідомлення 1024-бітове значення, яке усікається до необхідного розміру [18].

HAVAL – одно направлена хеш-функція. Для довільного вхідного повідомлення функція генерує хеш-значення, яке називається дайджестом повідомлення, яке може мати довжину 128, 160, 192, 224 або 256 біт. Кількість ітерацій — змінна, від 3 до 5. Кількість раундів на кожній ітерації – 32. Є модифікацією MD5.

Кессак – алгоритм хешування змінної розрядності. Алгоритм побудований за принципом криптографічної губки [9]. Вхідне повідомлення розбивається на блоки по 1024 біт [19].

N-Hash – криптографічна хеш-функція на основі циклічної функції FEAL. В даний час вважається небезпечною [20].

RIPEMD-128 – хеш-функція з розміром хешу 128-біт. Хешування RIPEMD-128 потребує достатньо багато часу для виконання і не гарантує високий рівень безпеки.

RIPEMD-160 – покращена версія RIPEMD, яка, в свою чергу, використовувала принципи MD4 і по продуктивності її можна порівняти з більш популярною SHA-1.

RIPEMD-256 – покращена версія RIPEMD-128 і призначена для додатків, які вимагають довшого хеша і не потребують більшої безпеки, ніж додатки рівня RIPEMD-128.

RIPEMD-320 – покращена версія RIPEMD-160, и призначена для додатків, які потребують більш довгого хешу и не потребують більшої безпеки, ніж додатки рівня RIPEMD-160.

Skein - алгоритм хешування змінної розрядності. Хеш-функція Skein виконана як універсальний криптографічний примітив, на основі блочного шифру Threefish, працюючого в режимі UBI-хешування [21].

Snefru – криптографічна односпрямована хеш-функція. Функція Snefru перетворює повідомлення довільної довжини в хеш-довжину [22].

Tiger – хеш-функція, призначена для особливо швидкого виконання на 64-розрядних комп'ютерах [23].

Проаналізуємо властивості найпоширеніших з цього списку алгоритмів (функцій) хешування та їх аналоги: MD5, HAVAL, SHA-256, Кессак.

2.2.1 Алгоритм MD5

Алгоритм MD5 є найбільш широко відомим і використовуваним алгоритмом хешування серед алгоритмів MD.

Алгоритм MD5 був запропонований в 1992 році, тобто всього через два роки після опублікування алгоритму MD4. Передумови розробки MD5 були такі [24]: при

розробці алгоритму MD4 основним критерієм була швидкодія даного алгоритму, його криптостійкість була дещо менш важлива. В результаті, вже в 1991 році була опублікована перша атака на усічений алгоритм MD4 [25].

Оскільки алгоритм MD4 вийшов виключно швидким, особливо на 32-бітних платформах, передбачалося, що саме тому він буде в центрі уваги криптоаналітиків, що, рано чи пізно, призведе до виявлення вразливостей і до успішних криптоаналітичних атак на даний алгоритм [26].

В результаті Рональд Рівест розробив алгоритм MD5 - дещо повільніший, ніж MD4, але маючий більш консервативний дизайн і враховуючий поточні результати і тенденції в криптоаналізі, що зумовило дещо вищу, порівняно з MD4, криптостійкість даного алгоритму [32].

Як і алгоритм MD4, MD5 виробляє 128-бітове хеш-значення для повідомлень перемінної довжини [33].

Перш за все, аналогічно алгоритму MD4, вхідне повідомлення розбивається на блоки по 512 бітів; останній з блоків завжди доповнюється до 512-бітового розміру таким чином:

КРОК 1 додається одиничний біт;

КРОК 2 потім - нульові біти до розміру 448 бітів; нульові біти не додаються, якщо після додавання одиничного біта блок має 448-бітовий розмір;

КРОК 3 до послідовності додається 64-бітове значення, рівне розміром в бітах вихідної послідовності вхідних даних до доповнення;

КРОК 4 якщо розмір повідомлення перевищує $2^{64} - 1$, то записується значення розміру по модулю 2^{64} [33].

Аналогічно алгоритму MD4, для зберігання проміжних значень після обробки кожного блоку повідомлення використовуються 32-бітові регістри A , B , C і D . Вектор ініціалізації алгоритму MD5 еквівалентний співвідношенню в MD4 [33].

512-бітові блоки доповненого повідомлення M_i по чергово обробляється таким чином:

КРОК 1 Блок вхідних даних представляється у вигляді 16 слів $M[0]...M[15]$, кожне з яких є 32-бітовим.

КРОК 2 Вміст регістрів $A...D$ копіюється в тимчасові змінні $a...d$.

КРОК 3 Виконуються 64 ітерації перетворень, в кожній з яких модифікуються змінні за участю одного із слів блоку повідомлення.

КРОК 4 Остання ітерація при обробці кожного блоку повідомлення відрізняється від інших [33].

Значення регістрів $A...D$ складаються по модулю 2^{32} з отриманими значеннями змінних $a...d$ відповідно.

Вихідне значення алгоритму, тобто 128-бітове хеш-значення повідомлення - результат конкатенації регістрів $A...D$ після обробки останнього блоку розширеного повідомлення.

Основні відмінності алгоритму MD5 від MD4 такі:

- в алгоритмі виконується 4 раунди замість трьох;
- змінені раундові функції; зокрема, більше не використовується функція, властивості якої викликали сумніви у криптологів ще в 1990 році;
- в кожен ітерацію додана додаткова операція додавання зі змінною b (саме змінна b модифікується в кожній попередній ітерації, за винятком першої) по модулю 2^{32} , яка прискорює вплив значень бітів вхідних даних на біти вихідного значення;
- в кожному раунді тепер використовуються унікальні модифікуючі константи;
- змінений (на менш упорядкований) порядок обробки вхідних слів блоку повідомлення, а також оптимізовані величини.

Алгоритм MD5 не виглядає більш складним в реалізації, ніж алгоритм MD4.

Всі зміни в алгоритмі є досить локальними, але в сукупності автору алгоритму

вдалося вирішити завдання його посилення за рахунок деякого погіршення швидкісних якостей алгоритму (згідно [26], MD4 приблизно в 1,5 рази швидше, ніж MD5).

Алгоритм MD5 широко використовувався донедавна і продовжує широко використовуватися, незважаючи на те, що існують відомі криптоаналітичні атаки проти нього.

В ході проведених досліджень вченими було виявлено цілу низку прогалин і вразливостей в цьому алгоритмі. Основною причиною слабкої захищеності MD5 є відносно легке знаходження колізій при шифруванні [7].

Під колізією розуміють можливість отримання однакового результату обчислень хеш-функції при різних вхідних значеннях.

2.2.2 Хеш-функція HAVAL

HAVAL - це односпрямована хеш-функція змінної довжини [27]. Вона є модифікацією MD5. HAVAL обробляє повідомлення блоками по 1024 біта, в два рази більшими, ніж у MD5. Використовується вісім 32-бітових змінних зчеплення, в два рази більше, ніж в MD5, і змінне число етапів, від трьох до п'яти (в кожному 16 дій). Функція може видавати хеш-значення довжиною 128, 160, 192, 224 або 256 бітів [27].

HAVAL замінює прості нелінійні функції MD5 на сильно нелінійні функції семи змінних, кожна з яких задовольняє суворому лавинному критерію. На кожному етапі використовується одна функція, але при кожній дії вхідні змінні переставляються різним чином. Використовується новий порядок повідомлення, і при кожному етапі (крім першого етапу) використовується своя константа, що додається. В алгоритмі також використовується два циклічних зсуви [35].

Змінна кількість етапів і змінна довжина видаваного значення означають, що існує 15 версій алгоритму [28]. Злам MD5, виконаний ден Боєром і Босселаєрсом, не застосовується до HAVAL через циклічний зсув функції [28].

У 2003 році Bart Van Rompay, Alex Biryukov та інші виявили колізію для 3-ітераційного HAVAL. Для знаходження цієї колізії було потрібно приблизно 2^{29} виконань функції стиснення H. У 2004 році китайські дослідники Wang Xiaoyun, Feng Dengguo, Lai Xuejia і Yu Hongbo оголосили про виявлену ними уразливості в 3-ітераційному HAVAL-128, що дозволяє за 27 HAVAL-обчислень знаходити колізії. У 2006 році група китайських вчених на чолі з Wang Xiaoyun і Yu Hongbo провели дві атаки на 4-ітераційний HAVAL, що зажадали 2^{43} і 2^{36} операцій хешування відповідно. Вони ж запропонували першу теоретичну атаку на 5-ітераційний HAVAL з числом операцій хешування, приблизно рівним 2^{123} [12].

2.2.3 Алгоритм безпечного хешування (Secure Hash Algorithm, SHA)

Принципи, що лежать в основі SHA аналогічні використаним професором Рональдом Л. Рівестом з MIT при проектуванні алгоритму короткого змісту повідомлення MD4. SHA розроблений за зразком згаданого алгоритму. SHA видає 160-бітове хеш-значення, більш довге, ніж у MD5 [36].

По-перше, повідомлення доповнюється, щоб його довжина була кратна 512 бітам. Використовується те же доповнення, що і в MD5: спочатку додається 1, а потім нулі так, щоб довжина отриманого повідомлення була на 64 біта менше числа, кратного 512, а потім додається 64-бітове представлення довжини оригінального повідомлення [36].

Ініціалізуються п'ять 32-бітових змінних (в MD5 використовується чотири змінних, але розглянутий алгоритм повинен видавати 160-бітове хеш-значення).

Потім починається головний цикл алгоритму. Він обробляє повідомлення 512-бітовими блоками і продовжується, поки не вичерпаються всі блоки повідомлення [36].

Головний цикл складається з чотирьох етапів по 20 операцій в кожному (у MD5 чотири етапи по 16 операцій в кожному). Кожна операція являє собою

нелінійну функцію над трьома з a , b , c , d і e , а потім виконується зсув і складання аналогічно MD5 [36].

SHA дуже схожий на MD4, але видає 160-бітове хеш-значення. Головною зміною є введення розширюючого перетворення і додавання виходу попереднього кроку в наступний з метою отримання більш швидкого лавинного ефекту. Рон Рівест опублікував цілі, переслідувані ним при проектуванні MD5, але розробники SHA цього не зробили [32]. Ось поліпшення, внесені Рівестом в MD5 щодо MD4, і їх порівняння з SHA:

- додався четвертий етап: у SHA також, однак у SHA на четвертому етапі використовується та ж функція f , що і на другому етапі;
- тепер в кожній дії використовується унікальна константа, що додається. SHA дотримується схеми MD4, повторно використовуючи константи для кожної групи їх 20 етапів;
- функція G на етапі 2 з була змінена, щоб зробити G менш симетричною. У SHA використовується версія функції з MD4;
- тепер кожна дія додається до результату попереднього етапу, це забезпечує більш швидкий лавинний ефект; ця зміна була внесена і в SHA; відмінність полягає в тому, що в SHA додана п'ята змінна; ця незначна зміна робить застосування злому MD5 ден Боєром і Босселаєрсом неможливим по відношенню до SHA;
- змінився порядок, в якому використовувалися підблоки повідомлення на етапах 2 і 3, щоб зробити шаблони менш схожими; SHA в цьому місці зовсім відрізняється, так як використовує циклічний код виправлення помилок;
- значення циклічного зсуву вліво на кожному етапі були наближено оптимізовані для прискорення лавинного ефекту - чотири зсуви, використовувані на кожному етапі, відрізняються від значень, використовуваних на інших етапах; SHA на кожному етапі використовує

постійне значення зсуву; це значення - взаємно просте чисте розміром слова, як і в MD4 [36].

Це призводить до наступного висновку: SHA - це MD4 з додаванням розширюючого перетворення, додаткового етапу і поліпшеним лавинним ефектом. MD5 - це MD4 з поліпшеним бітовим хешуванням, додатковим етапом і поліпшеним лавинним ефектом [25].

Відомості про успішні криптографічні зломи SHA відсутні. Так як ця односпрямована хеш-функція видає 160-хеш-значення, вона стійкіша до злому грубою силою (включаючи злом методом дня народження), ніж інші 128-бітові хеш-функції.

SHA-256 являє собою односпрямовану функцію для створення цифрових відбитків фіксованої довжини (256 біт, 32 байт) з вхідних даних розміром до 2,31 ексабайт (2^{64} біт) і є окремим випадком алгоритму з сімейства криптографічних алгоритмів SHA-2 (Secure Hash Algorithm Version 2) опублікованим АНБ США в 2002 році.

Хеш-функції сімейства SHA-2 побудовані на основі структури Меркле-Дамгарда.

Оригінал тексту після доповнення розбивається на блоки, кожен блок - на 16 слів. Алгоритм пропускає кожен блок повідомлення через цикл з 64 ітераціями. На кожній ітерації 2 слова перетворюються, функцію перетворення задають інші слова. Результати обробки кожного блоку складаються, сума є значенням хеш-функції. Так як ініціалізація внутрішнього стану проводиться результатом обробки попереднього блоку, то немає можливості обробляти блоки паралельно [25].

На поточний момент відомі методи для конструювання колізій до 31 ітерації. З огляду на алгоритмічної схожості SHA-2 з SHA-1 і наявності в останньої потенційних вразливостей прийнято рішення, що SHA-3 буде базуватися на зовсім іншому алгоритмі. 2 жовтня 2012 року NIST затвердив в якості SHA-3 алгоритм Кессак [8].

2.2.4 Алгоритм Кессак (SHA-3)

SHA-3 (Secure Hash Algorithm Version 3), також іменований Кессак (Кечак), являє собою односпрямовану функцію для створення цифрових відбитків обраної довжини (в стандарті прийняті 224, 256, 384 або 512 біт) з вхідних даних будь-якого розміру, розробленим групою авторів на чолі з Йоаном Дайменом в 2008 році і прийнятим в 2015 в якості нового стандарту FIPS. Алгоритм працює за допомогою функції перемішування зі стисненням до обраного розміру “криптографічною губкою”.

Оригінальний алгоритм Кессак має безліч параметрів, що налаштовуються (розмір блоку даних, розмір статку алгоритму, кількість раундів в функції f і інші) з метою забезпечення оптимального співвідношення криптостійкості та швидкодії для застосування на обраній платформі.

Версія алгоритму SHA-3, має кілька відмінностей від оригінального алгоритму Кессак:

- відкинуті повільні режими $s = 768$ і $s = 1024$;
- спрощений алгоритм заповнення;
- введені “функції з подовженим результатом” (XOF, Extendable Output Functions) SHAKE128 і SHAKE256, для чого хешує повідомлення стало необхідно доповнювати “суфіксом” з 2 або 4 біт, в залежності від типу функції.

Основою функції стиснення алгоритму є функція f , що виконує перемішування внутрішнього стану алгоритму. Стан A представляється у вигляді масиву 5×5 , елементами якого є 64-бітові слова, заповненні від початку нульовими бітами (розмір стану становить $5 \times 5 \times 64 = 1600$ бітів), ініціюються тимчасові масиви B ($5 * 5 * 64$ біт), C і D ($5 * 64$ біт). Функція f виконує 24 раунди, в кожному з яких виробляються операції з індексами масиву по модулю 5 і накладення операцією XOR раундової константи на слово $A[0, 0]$.

Перед виконанням функції стискання накладається операція XOR фрагментів вихідного повідомлення з фрагментами вихідного стану. Результат обробляється функцією f . Дане накладення в сукупності з функцією стискання, що виконуються для кожного блоку вхідних даних, являють собою “вбираючу” (absorbing) фазу криптографічної губки. Результируюче хеш-значення обчислюється в процесі виконання “вичавлюючої” (squeezing) фази криптографічної губки, основу якої також становить описана вище функція f .

Хоча алгоритм SHA-2 поки не зламаний, прийняття нового стандарту вже зараз готує “запасний аеродром” на випадок компрометації старого. Конкурс на новий алгоритм був оголошений в 2007 році після того, як був скомпрометований алгоритм SHA-1. Це вселило побоювання, що і подібний до нього сімейство SHA-2 не встоїть. За словами фахівців NIST, алгоритм Кессак вибраний завдяки простому і елегантному дизайну і набагато кращою, ніж у більшості конкурентів, продуктивності і легкості реалізації на різних типах пристроїв. Алгоритм принципово відрізняється від SHA-2, а отже уразливості, які можуть бути знайдені в старому стандарті, швидше за все не торкнуться нового.

2.3 Сучасні методи виявлення подібних за змістом документів

Пошук нечітких дублікатів дозволяє припустити, чи є два об'єкти частково однаковими чи ні. Під об'єктом можуть розумітися текстові файли та інші типи даних. Ми будемо працювати з текстом, але зрозумівши, як працює алгоритм, нам не складе труднощів перенести реалізацію на необхідні нам об'єкти.

Перед нами не стоїть завдання визначити абсолютне значення схожості об'єктів, а так само виділення в кожному з об'єктів подібних елементів. Нам необхідно лише припустити, чи є об'єкти майже дублікатами чи ні.

Одним з таких алгоритмів є “Алгоритм шинглів” (“шингл” англійською означає “лусочки”).

2.3.1 Алгоритм шинглів

Реалізація алгоритму складається з кількох етапів:

- КРОК 1 канонізація текстів;
- КРОК 2 розбиття тексту на шингли;
- КРОК 3 знаходження контрольних сум;
- КРОК 4 пошук однакових підпоследовностей.

В алгоритмі шинглів реалізовано порівняння контрольних сум текстів. Як відомо, контрольні суми статичних функцій дуже чутливі до змін [10].

При пошуку майже дубліката нас не цікавлять знаки пунктуації, сполучники, прийменники. Нас цікавлять тільки слова. Отже, ставиться завдання: очистити текст від непотрібних нам знаків і слів, які не несуть сенсу при порівнянні, це і називається “надати тексту канонічної форми”. Додатково можна, наприклад, кожен з іменників приводити до однини та називного відмінку. Для цього потрібно підключати морфологічні аналізатори.

Після того, як текст канонізовано, необхідно розбити кожен з них на підпоследовності - шингли. З виділених нами шинглів далі будуть виводитися контрольні суми.

Сам по собі алгоритм вимагає досить немало ресурсів, тому не завадить для нього створити механізм кешування, який буде особливо актуальним для порівняння наборів файлів, щоб не вираховувати контрольні суми кожного разу.

Так само, для збільшення продуктивності при обробці великих обсягів тексту можна порівнювати не всі отримані контрольні суми, а тільки ті, які, наприклад, діляться на 25, або будь-яке ціле число в межах від 10 до 40. Як показали тести, це дає значний приріст швидкості і не сильно зменшує точність, але тільки при обробці великих обсягів.

Існують модифіковані версії “Алгоритму шинглів” - “Алгоритм супершінглів” і “Алгоритму мегашинглів”.

2.3.2 Пошук схожих документів з MinHash

Алгоритм MinHash розроблений як продовження алгоритму Shingles. Він потребує менше часу на порівняння образів різних текстів і меншу кількість займаної пам'яті. Після обробки тексту алгоритмом шинглів алгоритм MinHash здійснює обчислення хеш-функцій і відбір їх мінімальних значень. Алгоритм MinHash використовує h хеш-функцій. Кожній хеш-функції ставиться у відповідність число – найменше її значення для всіх шинглів досліджуваного рядка. Отриманий в результаті вектор довжини h утворює сигнатуру або короткий числовий образ документа. На вхід алгоритму MinHash подається масив шинглів довжини N . Для порівняння будь-якого документа за допомогою MinHash його необхідно спочатку перетворити в множину елементів. Для текстів непогано підійде перетворення в множину ID слів. Також можна розбити тексти на N -грами. На виході алгоритм видає одновимірний масив (вектор) хеш-кодів довжини h , тобто сигнатуру або короткий числовий образ документа [11].

Алгоритм MinHash використовує однаковий набір хеш-функцій в заданому порядку для обробки всієї множини текстів, тому список хеш-функцій зазвичай є частиною реалізації алгоритму. Порівняння документів здійснюється за значеннями векторів `hashes1` і `hashes2`, побудованих алгоритмом MinHash з використанням одного і того ж складу хеш-функцій. Побудовані вектори `hashes1` і `hashes2` завжди мають однакову довжину, рівну h . Для виявлення дублікатів в алгоритм MinHash вбудовано 60 хеш-функцій.

Цей алгоритм відрізняється від інших своєю неймовірною швидкістю і здатен набагато пришвидшити пошук ймовірних дублікатів.

2.3.3 Пошук схожих документів з SimHash

SimHash є хеш-функцією, і її особливість полягає в тому, що чим більше схожих текстових рядків на вході, тим менша відстань Хеммінга в їх хешах (відстань Хеммінга - кількість позицій, на яких відповідні символи різні). Алгоритм

працює, розбиваючи текст на шматки та хешуючи кожен фрагмент вибраною функцією хешування [13]. Кожен хешований фрагмент представлений у вигляді бінарного вектора, і бітові значення перетворюються на +1 або -1 залежно від того, значення біта дорівнює 1 або 0. Для отримання SimHash ми складаємо бітові вектори. Нарешті, отримані біти встановлюються на 0, якщо сума негативна, інакше - 1 [12].

Алгоритм був розроблений одним із співробітників Google для порівняння великих наборів даних. Метод дозволяє створювати сигнатури для наборів даних. Порівняння сигнатур замість перевірки всього набору дає великий виграш у швидкості. Для створення сигнатури необхідно виконати такі дії:

КРОК 1 вибираємо розмір сигнатури;

створюємо масив V такого ж розміру, як кількість бітів у сигнатурі, і заповнюємо його нулями;

КРОК 2 вибираємо хеш-функцію, що дає результати розміру, аналогічно вибраному;

КРОК 3 для кожного набору, для кожного елементу в наборі розраховуємо значення хеш-функції;

КРОК 4 для кожного отриманого хешу в наборі, для кожного біта i :

- якщо біт = 1, додаємо 1 до $V[i]$;
- якщо біт = 0, віднімаємо 1 від $V[i]$;

КРОК 5 для кожного елементу i масиву V :

- якщо $V[i] > 0$, $V[i] = 1$;
- якщо $V[i] \leq 0$, $V[i] = 0$.

Отримані сигнатури матимуть вигляд двійкових кодів. Виконавши операцію побітового додавання по модулю 2 (XOR), отримаємо двійкове значення, яке є умовним значенням відмінності наборів. Завдяки цьому значенню можна вирахувати відсоток подібності наборів [19].

SimHash у порівнянні з MinHash дуже швидкий і потребує менше пам'яті. Але він накладає обмеження на те, наскільки різними повинні бути документи, щоб бути виявленими цим алгоритмом як дублікат. В основному він буде знаходити документи, які є майже однаковими.

MinHash використовує більше пам'яті та ресурсу процесора. Але він дозволяє охопити більш віддалену схожість, ніж з SimHash.

2.4 Розробка методу отримання цифрових відбитків документів з можливістю виявлення близьких дублікатів

Проаналізувавши сучасні алгоритми криптографічного хешування та дослідивши методи виявлення подібних за змістом документів, запропонуємо метод виведення цифрових відбитків творів, які будуть мати однаковий фіксований розмір та дадуть змогу виявляти не тільки дублікати, але й близькі за змістом твори.

Алгоритм MinHash дозволить знайти дублікат документа, а також виявить дуже схожі документи, але дасть змогу охопити більш віддалену схожість, ніж SimHash.

Алгоритм Кессак дозволить вивести хеш документу, захищений від розшифрування та колізій. Хоча результати SHA-256 ще не вдалося зламати, але виберемо алгоритм SHA-3 як інвестицію в майбутнє, оскільки він стійкіший за попередній. Крім того після хешування ми можемо отримати цифровий відбиток документу фіксованої довжини в 224, 256, 384 або 512 біт.

Пропонуємо об'єднати алгоритм MinHash, який виявляє близькі за змістом документи, та алгоритм Кессак для отримання цифрового відбитку результату функції MinHash, виконаної над документом. Оскільки Кессак - це одностороння функція, що для одних і тих же вхідних даних повертає один і той же хеш, вона поверне один і той же 224, 256, 384 або 512 бітовий відбиток сигнатури (короткого числового образу) документа. Це дозволить зберігати хеші документів у системі та

дасть можливість під час реєстрації документу виявити дублікат, тобто порушення авторських прав, та не дасть зареєструвати авторське право на такі документи.

2.5 Висновки по розділу

В даному розділі проаналізовано та вдосконалено метод отримання унікальних цифрових відбитків документів у вигляді хешів, які створюють колізію у разі хешування цим методом подібних документів. Це дозволить розробити систему, яка надаватиме можливості первинної реєстрації творів авторами, реєстрації передачі авторських прав між фізичними та юридичними особами та підтвердження (доведення) авторських прав чи прав власності на твори, однією із головних особливостей якої буде збереження цифрового відбитку твору замість збереження оригінального контенту файлу у вигляді хешу, що дозволить виявити плагіат чи близький за змістом твору разі повторної спроби реєстрації такого твору в системі.

3 ОПИС ПРОГРАМНОГО ТА ТЕХНІЧНОГО ЗАБЕЗПЕЧЕННЯ

3.1 Засоби розробки

Вибираючи засоби розробки інформаційної системи, розглянемо можливість використання засобів, що вільно розповсюджується та підтримуються товариствами. Open Source - програмне забезпечення з відкритими вихідними кодами. Вихідний код такого ПО доступний для перегляду і зміни. І це не просто технологічне визначення, що означає технічну доступність вихідних кодів програми, це філософське поняття, яке має на увазі, що вихідний код ПО повинен бути доступний всім, а гроші слід брати лише за конкретні послуги (впровадження, підтримку, навчання, доопрацювання).

У відкритого ПЗ є конкретні, відчутні переваги:

Передбачуваність. Користувач може забезпечити собі повну інформаційну безпеку системи. Доступні вихідні коди забезпечують можливість провести перевірку на наявність шпигунських або інших вставок.

Низька вартість. Користувач платить не за використання самих програм, а тільки за технічну підтримку, в тому випадку, якщо вона потрібна. Таким чином, виключається вартість ліцензійних відрахувань.

Повнота покриття. На сьогоднішній день світ продуктів Open Source включає в себе практично всі необхідні компоненти для автоматизованого управління всіма бізнес-процесами: серверні рішення, десктопні системи, спеціалізовані бізнес-додатки, що дозволяє гнучко підходити до вибору рішень з управління бізнес-процесами.

Швидкість реалізації доробок. Швидкість реакції на необхідність внесення змін в цілому вище, ніж при використанні закритих комерційних продуктів.

Повна відповідність потребам. Програмні продукти можна модифікувати і вдосконалити з метою найбільш повної відповідності специфічних потреб користувачів.

Незалежність. Знижується залежність від виробників програмних продуктів, від їх ліцензійної та цінової політики.

Головним і, мабуть, єдиним недоліком Open Source ПО є необхідність наявності кваліфікованого персоналу, який міг би успішно провести впровадження системи і здійснювати технічну підтримку, розробити керівництва користувачів і здійснювати доопрацювання. Іншими словами Open Source не дає гарантій того, що в ПЗ не буде помилок, що воно буде правильно впроваджено, що в процесі використання не відбудеться втрати даних і що завжди буде доступна служба підтримки (відзначимо, однак, що більшості цих гарантій не дає і комерційне програмне забезпечення).

Для усунення недоліків Open Source користувачі продуктів об'єднуються навколо кожного подібного проекту в співтовариства, які здійснюють взаємодопомогу. Крім цього виростають компанії, які на платній основі здійснюють повний спектр високоякісних послуг з цього продукту [20].

3.1.1 Інтегроване середовище розробки (IDE)

Visual Studio Code являється проектом з відкритим кодом та розповсюджується безкоштовно. Він поєднує простоту редактора вихідного коду з потужними інструментами для розробників, такими як доповнення коду за допомогою IntelliSense та відлагодження коду. Він забезпечує зручний та приємний цикл редагування-збирання-відлагодження без додаткових зусиль, що означає витрачання на це меншої кількості часу, ніж можна було б витратити на кожну функцію окремо, якби кожну необхідно було б виконувати за допомогою окремого засобу.

В основі програми Visual Studio Code є блискавично швидкий редактор вихідного коду, який ідеально підходить для щоденного використання. Завдяки підтримці сотень мов, VS Code допомагає бути продуктивним завдяки підсвічуванню синтаксису, відповідності дужок, автоматичному відступу, швидкій

вставці фрагментів тощо. Інтуїтивні комбінації клавіш, просте налаштування та відображення комбінацій клавіш для клавіатури, що надаються спільнотою, дозволяють легко редагувати код.

Код Visual Studio включає вбудовану підтримку семантики коду, навігацію та засобів рефакторингу коду. Visual Studio Code включає інтерактивний відладчик, за допомогою якого можна переходити через вихідний код, перевіряти змінні, переглядати стеки викликів і виконувати команди в консолі.

VS Code також інтегрується з інструментами побудови та створення сценаріїв для виконання звичайних завдань, що сприяє пришвидшенню повсякденного робочого процесу. VS Code має підтримку Git, тому у ньому можна керувати джерелами, не виходячи з редактора, включаючи перегляд змін.

3.1.2 Інструмент побудови діаграм

Lucidchart - це потужний інструмент діаграмування, який надає підприємствам та професійним особам веб-платформу, зручну для роботи в Інтернеті, завантажену безліччю функцій та можливостей, щоб допомогти їм у задоволенні їх потреб в розробці діаграм. Lucidchart, який використовується більш ніж 6 мільйонами користувачів у всьому світі, дозволяє зображати потоки (процеси) та будувати організаційні графіки, UML діаграми та багато інших.

Lucidchart - ідеальний інструмент діаграм для мозкового штурму та управління проектами. Інструмент також бездоганно працює з популярними веб-додатками та бізнес-системами, включаючи Google Apps. Lucidchart настільки інтуїтивно зрозумілий, що його використовують у багатьох галузях, включаючи інжиніринг, веб-дизайн та розробку, бізнес-сектор.

Створення діаграм для передачі ідей, представлення структури та організації, а також створення планів та карт проектів вже давно є ефективною практикою для незалежних фахівців для малого бізнесу для великих підприємств. За допомогою Lucidchart люди мають потужну та надійну платформу для виготовлення діаграм та

чартів, яка пропонує комфортне користування та просте навчання, гарантуючи, що вони отримують максимум від системи майже одразу з початку використання.

Що в основному пропонує Lucidchart - це платформу, яка дає змогу створювати привабливі та професійні схеми майже для всього, що потрібно для роботи - від ділових презентацій, схем організації робочого місця, до ілюстрацій виїзду з аварійного виходу. Користувачу не потрібно бути професійним дизайнером, щоб повністю використовувати систему. Lucidchart настільки зрозумілий, що замальовувати та обмінюватися професійними блок-схемами не складає труднощів.

Це популярне онлайн програмне забезпечення для діаграм адаптується до будь-яких робочих процесів, а це означає, що нам не потрібно змінювати спосіб роботи, щоб включити Lucidchart у нашу систему. Для зручності використання Lucidchart інтегрується з багатьма світовими широко використовуваними і не менш популярними бізнес-системами та програмами, такими як Google Apps, Google Drive, JIVE, Atlassian та багато інших. Окрім створення привабливих діаграм і діаграм потоків, їх експорт у різні стандартні формати файлів чи публікація їх у різних оглядачах документів в Інтернеті не є проблемою.

3.1.3 Інструмент для розробки дизайну та прототипу інтерфейсу користувача

Прототипування - важлива частина будь-якого процесу цифрового проектування. Це допомагає нам моделювати наші дизайни, досліджувати різний досвід користувачів та перевіряти поїздки та потоки користувачів у межах наших проектів.

Навколо існує багато інструментів веб-дизайну, але в останні роки інструменти для створення прототипів дійсно досягли високого рівня. За допомогою цих інструментів для складання прототипів можна вирішувати будь-які проблеми, пов'язані з дизайном, чи втіленням ідеальної технічної ідеї. З прототипом можливо протестувати інтерфейс, щоб мати змогу перевірити його зручність та зрозумілість.

Більшість цих інструментів мають цінник, але деякі надають вільний доступ до одного проекту або обмеженої кількості сторінок, що дозволяє нам знайти потрібну програму для свого проекту.

Proto.io - це неймовірний конкурент для створення реалістичних прототипів від грубих ідей аж до повноцінного дизайну. Інструмент також пропонує багато можливостей для проектів, включаючи створення детальної анімації та користувацької векторної анімації.

Це один із кращих інструментів, які допоможуть створити прототипи. З Proto.io запуск проекту стає простим. Замальовуючи лише грубу ідею, ми можемо зрозуміти недоліки відразу і почати використовувати заздалегідь налаштовані компоненти, щоб швидко розвивати свою концепцію.

Після того, як ми завершимо змальовувати ідеї, їх потім можна перетворити на більш відшліфовані та презентабельні макети для подальшої перевірки та роздумів. Наступний важливий крок після цього дає змогу налаштувати дизайн своїх прототипів, щоб вони були максимально відповідними кінцевому результату.

Proto.io - це веб-інструмент, який постачається з великою кількістю додатків та функцій, які включають імпорт файлів Sketch та Photoshop, VR-прототипування, бібліотеки інтерфейсу та навіть запис екрану для тестування. Перспектива створення детальної анімації насправді робить його чудовим інструментом для використання та ідеальним способом реалізувати свої ідеї максимально реалістично.

3.1.4 Система контролю версій вихідних кодів

Над проектами, як правило, паралельно працюють декілька розробників. Отже, потрібна система контролю версій, щоб не було конфліктів коду між розробниками.

Крім того, вимоги до проектів часто змінюються. Таким чином, система контролю версій дозволяє розробникам повернутися до більш старої версії коду.

Іноді кілька проектів, які виконуються паралельно, включають спільну кодову базу. У такому випадку дуже важлива концепція розгалуження.

Завдання, що описані вище, вирішує розподілена система контролю версій з відкритим кодом Git.

Код, який зберігається в Git, постійно змінюється, оскільки додається більше коду. Також багато розробників можуть паралельно додавати код. Тож система управління версіями допомагає в цьому, зберігаючи історію змін. Також Git надає такі функції, як розгалуження та злиття.

Робота з гілками є хорошою практикою, оскільки операція відділення дозволяє створити ще одну лінію розвитку. Ми можемо провадити процес розробки двома або більше різними способами. Наприклад, ми можемо створити гілку для роботи з основною версією системи та іншу гілку для роботи з експериментальними функціональними можливостями нашої системи.

Git має віддалене сховище, яке зберігається на сервері, і локальне сховище, яке зберігається в комп'ютері кожного розробника. Це означає, що код не просто зберігається на центральному сервері, але повна копія коду присутня на всіх комп'ютерах розробників. Git - це система управління розподіленою версією, оскільки код присутній на комп'ютері кожного розробника.

3.1.5 Хостинг віддаленого сховища Git

Для розробки інформаційної системи силами команди розробників необхідно забезпечити віддалене сховище на сервері, через який версії вихідного коду можуть контролюватись на всіх локальних оточеннях розробників. За допомогою централізованого віддаленого сховища розробники зможуть вести спільну роботу та обмінюватись змінами у вихідних кодах.

Привабливі умови пропонує хостинг BitBucket, оскільки він дозволяє безкоштовно створити приватний репозиторій для команди з п'яти розробників, на відміну від GitHub, у якого приватні репозиторії вимагають оплати.

BitBucket забезпечує нас наступними важливими функціями:

- контроль доступу, щоб обмежити доступ до вихідного коду;
- контроль робочого процесу для забезпечення розробки проекту чи роботи в команді;
- співпраця щодо перегляду коду;
- інтеграція з Trello для планування та відстеження розробки.

3.1.6 Інструмент планування та контролю процесу розробки

Trello — це простий, безкоштовний, гнучкий і візуальний сервіс для управління проектами й організації будь-чого. Інтерфейс Trello - це електронна дошка Kanban.

Дошка в Trello — це список списків, заповнених картками, якими можна користуватись одноосібно чи командою. Trello має все необхідне для того, щоб організувати проект будь-якого розміру.

Кожна картка призначена для роботи з одним завданням. В картці можна залишати коментарі, відправляти вкладення, створювати переліки, додавати мітки, дати завершення та багато іншого.

Можна запрошувати скільки завгодно людей до дошки абсолютно безкоштовно та додавати людей на картки для спільної роботи над завданнями. Усі учасники команди бачитимуть ту саму дошку й картину в цілому.

За допомогою системи сповіщень Trello учасники відразу можуть дізнатись, що сталося щось важливе. Вони отримують сповіщення де завгодно: в додатку, електронною поштою, в браузері на комп'ютері або через push-повідомлення на смартфоні. Сповіщення синхронізуються на всіх пристроях учасників.

Завдяки перелікам та датам завершення менеджери та розробники не пропустять важливі завдання, дати й пункти. З допомогою додатку «Календар» усе можна переглядати у вигляді зручного календаря всередині Trello.

Trello працює швидко та в режимі реального часу. Коли хтось переміщує картку, вона відразу ж рухається на дошках інших учасників команди, незалежно від того, які пристрої вони використовують. Вкладку можна весь день тримати відкритою без оновлення сторінки.

3.2 Архітектура програмного забезпечення

Проаналізувавши процеси і акторів у першому розділі роботи, можемо виділити такі сутності, які знайдуть відображення у класах інформаційної системи:

- Owner - власник права на твір або автор; він може володіти виключними або частковими правами на один і більше творів, що зареєстровані у системі;
- Work - твір; він може належати одному автору або декільком власникам;
- Certificate - свідоцтво про реєстрацію авторського права на твір; це свідоцтво буде підтверджувати реєстрацію одного твору від одного чи декількох авторів;
- Transfer - договір про передання прав власності на твір; за цим договором може бути передано виключне або часткове право власності одному або декільком правовласникам.

Між зазначеними класами необхідно встановити такі зв'язки:

- Owner-Transfer - багато до багатьох;
- Work-Certificate - один до багатьох;
- Work-Transfer - багато до багатьох.

Зобразимо класи, їх атрибути та відношення між ними на діаграмі класів (рисунок 3.1).

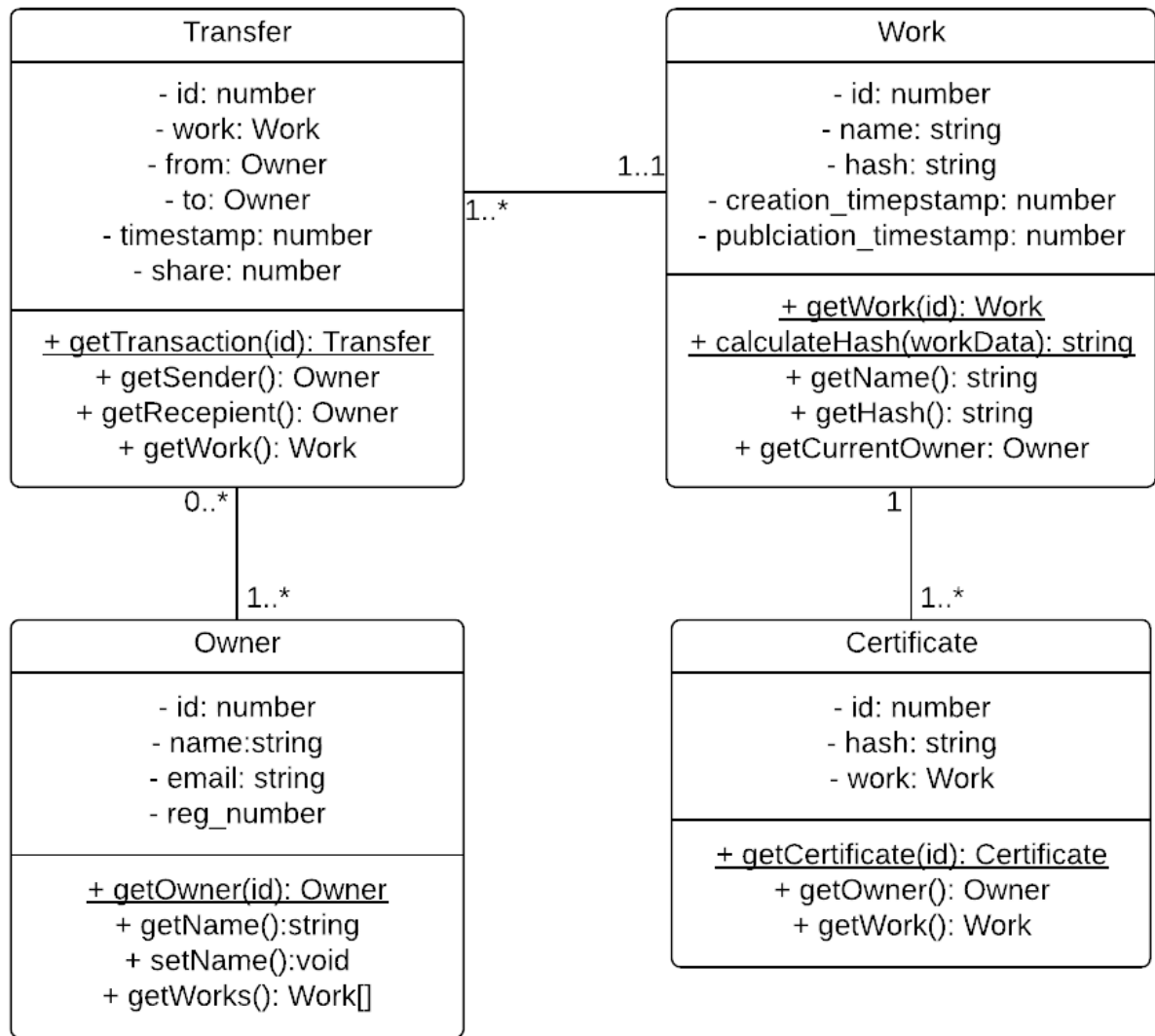


Рисунок 3.1 - Діаграма класів інформаційної системи реєстрації та підтвердження авторських прав

3.3 Інструкція користувача

Для того, щоб скористатися можливостями системи із реєстрації творів та права на них, передачі прав, користувачу необхідно зареєструватись або авторизуватись у системі. Якщо користувачу необхідно лише віднайти підтвердження права власності на твір певної фізичної чи юридичної особи, то це можна зробити без авторизації (як гість).

3.3.1 Реєстрація твору та авторського права на нього

Щоб виконати дію реєстрації твору та авторського права на нього потрібно в головному меню інтерфейсу користувача вибрати пункт “Зареєструвати твір”.

У розділі “Зареєструвати твір” потрібно заповнити одну з реєстраційних форм. Одна з форм призначена для реєстрації твору автором - фізичною особою, інша - юридичною особою.

Обов’язковими полями форми для фізичних осіб є:

- прізвище;
- ім’я;
- по батькові;
- ідентифікаційний податковий номер (або серія та номер паспорта, якщо ПІН немає);
- дата створення твору;
- файл - копія твору;
- фраза-пароль або публічний ключ.

Обов’язковими полями форми для юридичних осіб є:

- назва юридичної особи;
- реєстраційний номер;
- прізвище довіреної особи;
- ім’я довіреної особи;
- по батькові довіреної особи;
- дата та номер довіреності, виданої юридичною особою довірений особі для виконання операцій від її імені;
- дата створення твору;
- файл - копія твору;
- фраза-пароль або публічний ключ.

Не обов’язковими для заповнення в обох формах є поля:

- дата оприлюднення твору;

- місце оприлюднення твору.

Натиснувши на кнопку “Зареєструвати” дані будуть відправлені на сервер для перевірки та зберігання, у разі якщо такого твору не було зареєстровано у системі раніше.

Важливо! Зміст файлу не буде передано на зберігання у системі. Натомість зберігатися буде його цифровий відбиток, який не дасть змогу повторно зареєструвати твір від імені іншої особи або зареєструвати близький за змістом твір, що являється дублікатом зі змінами.

У разі успішної перевірки на відсутність дублікатів у системі та збереження цифрового відбитку твору разом з даними про його автора (правовласника), користувачу буде запропоновано здійснити оплату послуги онлайн. Після успішної оплати користувач отримує цифрове свідоцтво про реєстрацію з хешем його твору. Дане свідоцтво буде публічно доступне для перегляду у системі по посиланню на нього за допомогою унікального ідентифікатора. За бажанням його можна роздрукувати.

3.3.2 Реєстрація передачі авторського права на твір

Щоб виконати дію реєстрації передачі авторського права на твір потрібно в головному меню інтерфейсу користувача вибрати пункт “Передати твір”. Там користувачу буде запропоновано заповнити форму необхідними для виконання операції даними. Форма для передачі прав фізичній особі (особам) буде складатись з наступних полів:

- файл - копія твору або його хеш;
- фраза-пароль (якщо не використовуються криптографічні ключі);
- публічний ключ(-і), що належить(-ать) приймачу(-ам) прав;
- прізвище приймача прав;
- ім'я приймача прав;
- по батькові приймача прав;

- індивідуальний податковий номер приймача прав (або серія та номер паспорта).

Якщо необхідно передати авторське право декільком особам, то в інтерфейсі користувача є кнопка “Додати приймача”, за допомогою якої можна додати поля для передачі даних про всіх осіб, яким передається право.

Для передачі прав власності юридичній особі (особам), потрібно заповнити наступні поля:

- файл - копія твору або його хеш;
- фраза-пароль (якщо не використовуються криптографічні ключі);
- публічний ключ(-і), що належить(-ать) приймачу(-ам) прав;
- назва юридичної особи - приймача прав;
- реєстраційний номер юридичної особи - приймача прав;
- номер та дата договору про передачу прав на твір;
- дата та номер довіреності, виданої юридичною особою довіренту для виконання операцій від її імені.

Якщо необхідно передати авторське право декільком юридичним особам, то під формою розміщена кнопка “Додати приймача”, за допомогою якої можна додати поля для передачі даних про всіх юридичних осіб, яким передається право власності на твір.

Натиснувши на кнопку “Передати права” система перевірить правильність введення фрази-паролю або відповідність приватного ключа користувача публічному, що зареєстрований разом із твором у системі. Якщо перевірка пройде успішно, то користувач отримає запит на оплату послуги онлайн.

Після оплати твір буде перереєстровано за новим(и) поточним(и) власником(-ами). Передати права власності на твір або довести це право нові власники зможуть за допомогою фрази-паролю або приватних ключів, що відповідають публічним ключам, зареєстрованих при передачі.

При успішній реєстрації передачі прав новим власникам буде видано нові сертифікати про право власності на твір. Дані свідоцтва будуть публічно доступні для перегляду у системі по посиланням на них за допомогою їх унікальних ідентифікаторів. За бажанням їх можна роздрукувати.

Недійсні свідоцтва будуть анульовані, про що буде повідомлено при спробі їх перегляду за унікальними ідентифікаторами.

3.3.3 Підтвердження авторського права на твір

Щоб здійснити підтвердження авторського права на твір потрібно в головному меню інтерфейсу користувача вибрати пункт “Знайти твір”.

Для знаходження твору потрібно передати системі лише його електронну копію, або цифровий відбиток (хеш). Система надасть інформацію про поточного(-их) власника(-ів) за замовчуванням. У інтерфейсі користувача наявна опція запитати у системи ланцюжок передачі прав від автора(-ів) до поточного(-их) правовласника(-ів).

Для підтвердження даних, що надає система, та для захисту від будь-яких маніпуляцій даними власність над твором потрібно додатково підтвердити за допомогою фрази-пароллю або наявністю приватного ключа, що відповідає публічному, який зберігається у списку публічних ключів поточних власників авторських прав на шуканий твір.

3.4 Опис технічного забезпечення

Система реєстрації та підтвердження авторських прав передбачає роботу в глобальній мережі Інтернет та розгортання на віртуальному веб-сервері хмарного провайдера. Основними компонентами серверної частини являються веб-сервер, що обробляє запити від клієнтського програмного забезпечення, веде журнал подій та працює з сервером керування базою даних. Власне сервер керування базою даних є другим компонентом серверної частини розгортання системи (рисунок 3.5).

Клієнтське програмне забезпечення може бути будь-якого виду. Єдиною умовою є підтримка протоколів HTTP (Hyper Text Transfer Protocol) та HTTPS (HTTP Secure) та взаємодія з сервером з використанням підходу REST (Representational State Transfer). На початковому етапі розробки системи на стороні клієнта планується інтерфейс у вигляді веб-додатку, що буде доставлений користувачу за допомогою веб-браузера.

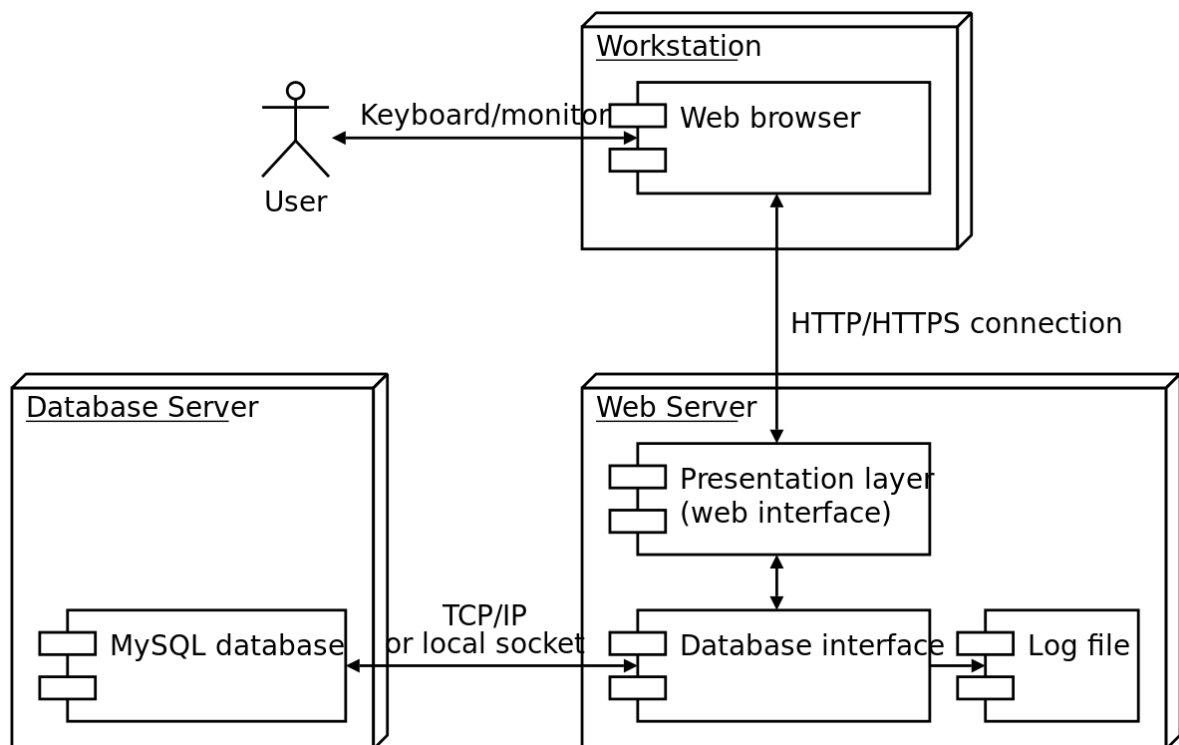


Рисунок 3.5 - Діаграма розгортання інформаційної системи реєстрації та підтвердження авторських прав

3.5 Висновки по розділу

Систему реєстрації та підтвердження авторських прав можливо розробити за допомогою відкритих та безкоштовних засобів розробки та управління проектами. Розгортання системи має класичну клієнт-серверну архітектуру та доставляється клієнтам як веб-додаток засобами мережі Інтернет. Початковий вигляд архітектури системи не є складним, тому для її імплементації не потрібно наймати спеціалістів

найвищого класу, що спростить пошук необхідних кадрів та зменшить витрати на їх утримання.

4 РОЗРОБКА СТАРТАП-ПРОЕКТУ

4.1 Опис ідеї проекту

Проаналізуємо зміст ідеї, що пропонується, можливі напрямки застосування, основні вигоди, що може отримати користувач товару за кожним напрямком застосування. Після аналізу ідеї проекту порівняємо наш проект з існуючими аналогами та заміниками, якщо такі наявні на ринку.

4.1.1 Зміст ідеї, напрямки та вигоди застосування

Подано цілісне уявлення про зміст ідеї та можливі базові потенційні ринки, в межах яких знайдемо групи потенційних клієнтів (табл. 4.1).

Таблиця 4.1 - Опис ідеї стартап-проекту

Зміст ідеї	Напрямки застосування	Вигоди для користувача
1	2	3
Надання можливості первинної реєстрації творів авторами, реєстрації передачі авторських прав між фізичними та юридичними особами та підтвердження (доведення) авторських прав чи прав власності на твори під час розгляду спорів у суді	Реєстрація та підтвердження немайнових прав на твір: авторства, права на публікацію, права на недоторканність твору і збереження його цілісності	1) Можливість реєстрації твору та підтвердження авторства без відвідування офісу спеціалізованої установи (організації) 2) Економія часу завдяки автоматизації процесів реєстрації та підтвердження авторства

Продовження таблиці 4.1

1	2	3
	Передача виключних чи часткових майнових прав на твір: права на відтворення, права на отримання винагороди за використання твору, права на розповсюдження	1) Можливість реєстрації передачі майнових прав без відвідування офісу спеціалізованої установи (організації) 2) Економія часу завдяки автоматизації процесу реєстрації
	Доведення майнових прав на твір шляхом відтворення усього ланцюжку передавання прав	1) Можливість доведення майнових прав без відвідування офісу спеціалізованої установи (організації) 2) Економія часу та коштів завдяки автоматизації процесу збору та відтворення повного ланцюга передачі майнових прав з урахуванням часток

Ідея автоматизації реєстрації та підтвердження авторських прав може бути використана у приватному та державному секторах. Особливо цікавою вона буде для виробників цифрових творів, наприклад програмного забезпечення або баз даних, хоча її можна застосувати до будь-яких художніх та нехудожніх творів.

4.1.2 Визначення сильних, слабких та нейтральних характеристик ідеї проекту

Дослідимо потенційні техніко-економічні переваги ідеї порівняно із пропозиціями конкурентів (табл. 4.2). Для цього виконаємо наступні завдання:

- визначимо перелік техніко-економічних властивостей та характеристик ідеї;
- визначимо попереднє коло конкурентів (проектів-конкурентів) або товарів-замінників чи товарів-аналогів, що вже існують на ринку;
- проведемо збір інформації щодо значень техніко-економічних показників для ідеї власного проекту та проектів-конкурентів відповідно до визначеного вище переліку;
- проведемо порівняльний аналіз показників: для власної ідеї визначаються показники, що мають гірші (слабші) значення, аналогічні (нейтральні) значення, кращі (сильні) значення.

Таблиця 4.2 - Визначення сильних, слабких та нейтральних характеристик ідеї проекту

№ п/п	Техніко-економічні характеристики ідеї	(потенційні) товари/концепції конку - № рентів		W (слабка сторона)	N (нейтрал ьна сторона)	S (сильна сторона)
		мій проект	державна система			
1	2	3	4	5	6	7
1	Собівартість надання послуг	підтримка серверної інфраструк тури та системи, захист інформації	утримання держслужбо вців, підтримка реєстру, захист інформації			+
2	Вартість надання послуг	підписка або разова плата	держмито	+		
3	Строки надання послуг	декілька хвилин	2-3 місяці			+

Продовження таблиці 4.2

1	2	3	4	5	6	7
4	Повнота наданої послуги	повний ланцюг передачі	часткова інформація з різних джерел, підняття всіх договорів			+

Основним невідновлюваним ресурсом у бізнесі є час. Як бачимо, ідея створення системи реєстрації та підтвердження захисту авторських прав має декілька потужних переваг, основною з яких є значна економія часу.

4.2 Технологічний аудит ідеї проекту

У даному розділі здійснимо аудит технології, за допомогою якої можна реалізувати ідею проекту. Визначимо технологічну здійсненність ідеї проекту (табл. 4.3). Проаналізуємо наступні складові:

- за якою технологією буде виготовлено товар згідно ідеї проекту;
- чи існують такі технології, чи їх потрібно розробити/доробити;
- чи доступні такі технології авторам проекту.

Таблиця 4.3 - Технологічна здійсненність ідеї проекту

№ п/п	Ідея проекту	Технології її реалізації	Наявність технологій	Доступність технологій
1	2	3	4	5
1	Програмне забезпечення	Програмне забезпечення як сервіс	Наявні	Доступні

Продовження таблиці 4.3

1	2	3	4	5
2	Програмне забезпечення	Локальне програмне забезпечення	Наявні	Доступні, але потребують значних вкладень для підтримки
3	Програмне забезпечення	Розподілена обчислювальна система без центрального вузла управління	Наявні	Доступні, але потребують участі більшої кількості користувачів, для забезпечення збереження цілісності даних
Обрана технологія реалізації ідеї проекту: Програмне забезпечення як сервіс				

За результатами аналізу технологічності здійснення ідеї проекту, можемо зробити висновок, що реалізація проекту можлива. Обрана технологія “Програмне забезпечення як сервіс” (Software-as-a-Service, SaaS) надасть можливість зосередитися на створенні та впровадженні проекту, а провайдер хмарних рішень подбає про цілодобову доступність програмного забезпечення та швидку доставку даних користувачу.

4.3 Аналіз ринкових можливостей запуску стартап-проекту

Проаналізуємо ринкові можливості запуску стартап-проекту. Для цього визначимо ринкові можливості, які можна використати під час ринкового впровадження проекту, та ринкових загроз, які можуть перешкодити реалізації проекту. Це дозволить спланувати напрями розвитку проекту з урахуванням стану ринкового середовища, потреб потенційних клієнтів та пропозицій проектів-конкурентів.

4.3.1 Аналіз попиту

На початку аналізу ринкових можливостей запуску стартап-проекту проведемо аналіз попиту, а саме: наявність попиту, обсяг, динаміку розвитку ринку (табл. 4.4).

Таблиця 4.4 - Попередня характеристика потенційного ринку стартап-проекту

№ п/п	Показники стану ринку (найменування)	Характеристика
1	Кількість головних гравців, од	1
2	Загальний обсяг продаж, грн/ум.од.	не відомо
3	Динаміка росту (якісна оцінка)	постійно зростає
4	Наявність обмежень для входу (вказати характер обмежень)	існує один великий гравець, який диктує умови в даному сегменті (держава)
5	Специфічні вимоги до стандартизації та сертифікації	немає
6	Середня норма рентабельності галузі (або ринку), %	> 100%

Станом на листопад 2019 року банківський відсоток на вклад становить 11-17%. Завдяки особливості розповсюдження програмного забезпечення як сервіс у наш час, рентабельність інформаційних систем становить більше ніж 100%. Отже, це більше ніж у п'ять разів більше, ніж можна отримати з банківського депозиту. Дивлячись на динаміку росту попиту на ринку, є сенс вкласти кошти в розробку даного проекту.

4.3.2 Потенційні групи клієнтів

Наступним кроком визначимо потенційні групи клієнтів, їх характеристики, та сформуємо орієнтовний перелік вимог до товару для кожної групи (табл. 4.5).

Таблиця 4.5 - Характеристика потенційних клієнтів стартап-проекту

Потреба, що формує ринок	Цільова аудиторія (цільові сегменти ринку)	Відмінності у поведінці різних потенційних груп клієнтів	Вимоги споживачів до послуг
<p>1. Швидко і надійно підтверджувати своє авторство на твір</p> <p>2. Швидко і надійно підтверджувати свої часткові або виключні майнові права</p> <p>3. Без зайвих зусиль доводити свої права на твір у суді</p>	<p>1. Особи-автори (письменники, поети, художники, композитори)</p> <p>2. Звукозаписувальні студії</p> <p>3. Кінематографічні компанії</p> <p>4. Виробники програмного забезпечення</p> <p>5. Розробники та власники баз даних та знань</p>	<p>1. Користування сервісом поодинокими авторами</p> <p>2. Користування послугами доведення виключних майнових прав юридичними особами</p> <p>3. Користування послугами доведення часткових майнових прав юридичними особами</p>	<p>1. Швидке та недороге обслуговування</p> <p>2. Зручне і зрозуміле користування сервісом без спеціального навчання</p> <p>3. Надійне збереження даних про належність творів</p> <p>4. Захист від реєстрації плагіату та майже дубльованих творів</p>

Характеристика потенційних клієнтів стартап-проекту показала, що найбільшою цільовою аудиторією є компанії, що займаються витворами мистецтва. Трохи меншою аудиторією є компанії, що виробляють програмне забезпечення та моделюють візуальні інтерфейси.

4.3.3 Аналіз ринкового середовища

Визначивши потенційні групи клієнтів, проведемо аналіз ринкового середовища. У таблиці 4.6 наведено фактори, що сприяють ринковому впровадженню проекту.

Таблиця 4.6 - Фактори загроз

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії
1	Законодавство	Невизнання недержавної організації захисту авторських прав державними інститутами	Подання до Верховної Ради законопроекту про захист авторських прав недержавними організаціями
2	Державна система захисту авторських прав	Створення державою автоматизованої системи реєстрації та підтвердження авторських прав	Впровадження конкурентних послуг та можливостей Проведення рекламної кампанії, направлену на зміцнення бренду
3	Поява прямих конкурентів	Вихід на ринок компаній, що пропонують схожі або тотожні послуги	Впровадження конкурентних послуг та можливостей Розширення спектру послуг, що можуть бути суміжними з основними Проведення рекламної кампанії, направлену на зміцнення бренду Посилення прямої роботи з клієнтами
4	Ліцензування	Необхідність ліцензування діяльності державними органами	Проходження сертифікації та отримання ліцензії на провадження діяльності щодо захисту авторських прав
5	Несприятлива ситуація у країні	Девальвація валюти, автоматичне підвищення ціни у національній валюті	Розробка та введення додаткових стимулів для покупців Зміна напрямку розвитку проекту до більш рентабельного

Оскільки головним завданням систем реєстрації та підтвердження авторських прав є доведення власних прав на твір у суді, то дуже важливим є визнання цієї системи судовою владою та іншими державними інститутами. Це є найбільшою загрозою дії проекту на теренах держави.

Далі перелічимо фактори, що перешкоджають ринковому впровадженню проекту (табл. 4.7).

Таблиця 4.7 - Фактори можливостей

№ п/п	Фактор	Зміст можливості	Можлива реакція компанії
1	Вихід на міжнародний ринок	Послуга не залежить від державного регулювання і має визнання державними інститутами інших країн	Збільшення території впливу та кількості постійних клієнтів Адаптація до умов міжнародного ринку та специфіки ринків інших країн
2	Вільний ринок	Відсутність конкурентів у секторі	Встановлення власних правил на внутрішньому ринку
3	Органічний розвиток	Нарощування функціоналу згідно потреб клієнтів.	Покращення сервісу та розширення списку наявних послуг

Найбільш позитивним фактором для компанії, у яку може бути перетворено стартап-проект згодом, може бути вихід на міжнародний ринок, що сприятиме ще більшому її росту та розвитку.

4.3.4 Аналіз пропозиції

Наступним етапом аналізу буде визначення загальних рис конкуренції на ринку (табл. 4.8).

Таблиця 4.8 - Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)
1	2	3

Продовження таблиці 4.8

1	2	3
1. Монополістична конкуренція	Складно завоювати долю на ринку через монополію державних інститутів, а також вийти на міжнародний рівень	Дії, направлені на зміцнення бренду, та продумана реклама сильних сторін проекту
2. Глобальний рівень конкурентної боротьби	Наявність конкурентів у відповідному ринковому сегменті інших країн	Подолання конкуренції на українському ринку та налагодження зв'язків з закордонними партнерами
3. Внутрішньогалузева	Надання подібних чи тих самих послуг в галузі захисту авторських прав	Фокусування на сильних сторонах проекту та розвиток унікальних послуг
4. Товарно-видова конкуренція	Конкуренція між різновидами послуг	Надання якісних, зручних та порівняно недорогих послуг
5. Ціновий характер	Конкуренція в ціні за надання послуг	Надання якісних, зручних та порівняно недорогих послуг
6. Марочний характер	Споживачів цікавить перевірений та відомий товар	Зміцнення бренду

Ступеневий аналіз конкуренції ринку показав, що не дивлячись на високу конкуренції з боку держави, все рівно є шанси здобути першість в Україні за рахунок дуже значного скорочення строків надання послуг та їх зручності, що в свою чергу призведе до збільшення кількості потенційних користувачів.

4.3.5 Детальний аналіз умов конкуренції в галузі

Після аналізу конкуренції проведемо більш детальний аналіз умов конкуренції в галузі за моделлю 5 сил М. Портера (табл. 4.9).

Таблиця 4.9 - Аналіз конкуренції в галузі за М. Портером

Складові аналізу	Прямі конкуренти в галузі	Потенційні конкуренти	Постачальники	Клієнти	Товари-замінники
	держава	рівень відомості бренду, спектр надаваних послуг	тарифи з надання послуг хмарними провайдерами	кількість реєстрацій, попит	ціна, лояльність споживачів
Висновки:	Держава надає послуги протягом декількох місяців, розглядає заяви протягом місяця	Можливості виходу на ринок такі ж як і у нашого проекту, строки виходу на ринок залежать від списку послуг, які планує надавати конкурент	Вартість послуг хмарних провайдер невелика в порівнянні з запланованими доходами, тому вони не впливають на нашу присутність на ринку	Клієнти диктують свої умови на ринку. Переважно це відноситься до якості та швидкості надання послуг, а також до збереження персональних даних та захисту авторства	Потенційні товари-замінники обмежують встановлення високих цін на послуги

За результатами аналізу таблиці можемо зробити висновок, що прямим конкурентом на території України є сама держава, але вона надає послуги із реєстрації та підтвердження авторських прав дуже повільно, що значно полегшує вихід на ринок України за 1 рік.

4.3.6 Фактори конкурентоспроможності

На основі аналізу конкуренції, проведеного в 4.3.5, а також із урахуванням характеристик ідеї проекту в 4.1.2, вимог споживачів до товару в 4.3.2 та факторів

маркетингового середовища в 4.3.3 та 4.3.4 визначимо та обґрунтуємо перелік факторів конкурентоспроможності (табл. 4.10).

Таблиця 4.10 - Обґрунтування факторів конкурентоспроможності

№ п/п	Фактор конкурентоспроможності	Обґрунтування (наведення чинників, що роблять фактор для порівняння конкурентних проектів значущим)
1	Попит внутрішнього ринку	Чим більше клієнтів користуються послугами, тим більший дохід
2	Вартість	Конкурентна помірна ціна
3	Простота експлуатації	Зрозумілий інтерфейс та виконання необхідних дій у кілька простих кроків робить систему привабливою для клієнтів
4	Терміни надання послуг	Обслуговування користувачів у найкоротший можливий термін

Обґрунтування факторів конкурентоспроможності показав, що існує фактор необхідності у створенні системи із зрозумілим інтерфейсом та найкоротшими строками реєстрації, передачі чи підтвердження авторських прав. Враховуючи те, що вітчизняний ринок у даній сфері майже вільний, попит на товар буде великим.

4.3.7 Аналіз сильних та слабких сторін стартап-проекту

За визначеними факторами конкурентоспроможності в 4.3.6 проведемо аналіз сильних та слабких сторін стартап-проекту (табл. 4.11).

Таблиця 4.11 - Порівняльний аналіз сильних та слабких сторін нашої системи

№ п/п	Фактор конкурентоспроможності	Бали 1-20	Рейтинг послуг - конкурентів у порівнянні з державними						
			-3	-2	-1	0	+1	+2	+3
1	2	3	4	5	6	7	8	9	10

Продовження таблиці 4.11

1	2	3	4	5	6	7	8	9	10
1	Попит внутрішнього ринку	15				+			
2	Вартість	10			+				
3	Простота експлуатації	16					+		
4	Терміни надання послуг	16						+	

Порівняльний аналіз сильних та слабких сторін “Системи реєстрації та підтвердження авторських прав” показав, що фактор конкурентоспроможності маркетингової пропозиції нашого проекту трохи сильніший ніж у держави. До сильних факторів конкурентоспроможності можна віднести простоту експлуатації та короткі терміни надання послуг.

4.3.8 SWOT - аналіз стартап - проекту

Фінальним етапом ринкового аналізу можливостей впровадження проекту є складання SWOT-аналізу (матриці аналізу сильних (Strength) та слабких (Weak) сторін, загроз (Troubles) та можливостей (Opportunities) на основі виділених ринкових загроз та можливостей, та сильних і слабких сторін. Результати наведені у таблиці 4.12.

Таблиця 4.12 - SWOT- аналіз стартап-проекту

<p>Сильні сторони:</p> <ul style="list-style-type: none"> - надійність та простота експлуатації - зрозумілі відео- та текстові інструкції з використання - автоматизація процесу перевірки документів - автоматизація оплати послуг - можливість надання послуг на міжнародному ринку - найкоротші терміни надання послуг та інформації - постійна підтримка клієнтів 	<p>Слабкі сторони:</p> <ul style="list-style-type: none"> - потрібно визнання або ліцензування від державних органів
<p>Можливості:</p> <ul style="list-style-type: none"> - вільний ринок, відсутні аналоги на внутрішньому ринку - вихід на міжнародний ринок для надання послуг в інших країнах 	<p>Загрози:</p> <ul style="list-style-type: none"> - поява конкурентів на внутрішньому ринку; - збереження державою монополії на надання послуг; - невизнання системи державними органами інших країн; - нестабільна ситуація у країні, знецінення валюти, припинення отримання інвестицій; - зменшення попиту при насиченні ринку.

Сильними сторонами проекту є простота в експлуатації, постійна підтримка клієнтів, найкоротші можливі терміни надання послуг за рахунок автоматизації. Можливо для визнання нашої системи державними інститутами знадобиться ліцензування або впровадження відповідного законопроекту, проте проект має можливості як на внутрішньому ринку через відсутність інших конкурентів, крім держави, а також на міжнародному як альтернатива вже існуючим там системам.

4.3.9 Альтернативи ринкової поведінки

На основі SWOT-аналізу розробимо альтернативи ринкової поведінки (перелік заходів) для виведення стартап-проекту на ринок та орієнтовний оптимальний час їх ринкової реалізації. Визначені альтернативи проаналізуємо з точки зору строків та ймовірності отримання ресурсів (табл. 4.13).

Таблиця 4.13 - Альтернативи ринкового впровадження стартап-проекту

№ п/п	Альтернатива (орієнтовний комплекс заходів) ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації
1	Запуск проекту в Україні	80%	9 місяців
2	Запуск проекту за кордоном	80%	9 місяців

Як бачимо з таблиці 4.13 проект має приблизно однакові шанси на запуск як в Україні так і за її межами, причому від цього не залежатимуть терміни, необхідні на розробку, розгортання та початок роботи системи.

4.4 Розроблення маркетингової програми стартап-проекту

4.4.1 Визначення стратегії охоплення ринку

Розроблення ринкової стратегії першим кроком передбачає визначення стратегії охоплення ринку. Опис цільових груп потенційних споживачів наведено у таблиці 4.14.

Таблиця 4.14 - Вибір цільових груп потенційних споживачів

№ п/п	Опис профілю цільової групи потенційних клієнтів	Готовність споживачів сприйняти продукт	Орієнтовний попит в межах цільової групи (сегменту)	Інтенсивність конкуренції в сегменті	Простота входу у сегмент
1	Художня сфера (автори)	готові	високий	відносно невелика	відносно легко
2	Наукова сфера	готові	високий	відносно невелика	відносно легко
3	Дослідницька сфера	готові	високий	відносно невелика	відносно легко
4	Освітницька сфера	готові	середній	відносно невелика	відносно легко
5	Сфера інформаційних технологій	готові	високий	відносно невелика	відносно легко
Які цільові групи обрано: автори, науковці, дослідники та винахідники, розробники інформаційних систем та програмного забезпечення					

За цільові групи потенційних споживачів обрано авторів, науковців, дослідників та винахідників, розробників інформаційних систем та програмного забезпечення, оскільки в них при високому попиті немає конкуренції, якщо не враховувати надання послуг державою.

4.4.2 Базова стратегія розвитку

Сформуємо базову стратегію розвитку для нашого стартап-проекту. Результати наведено у таблиці 4.15.

Таблиця 4.15 - Визначення базової стратегії розвитку

№ п/п	Обрана альтернатива розвитку проекту	Стратегія охоплення ринку	Ключові конкурентоспроможні позиції відповідно до обраної альтернативи	Базова стратегія розвитку
1	Розроблення онлайн-сервісу реєстрації та підтвердження авторських прав	Стратегія сегментної спеціалізації	Довіра до бренду, відповідна ціна, спрощення умов надання послуг, зменшення строків надання послуг	Стратегія диференціації

За базову стратегію розвитку обрано стратегію диференціації. Вона передбачає надання товару важливих з точки зору споживача відмінних властивостей, які роблять товар відмінним від товарів конкурентів. Реалізація цієї стратегії вимагає, як правило, більш високих витрат. Проте успішна диференціація дозволяє компанії домогтись більшої рентабельності за рахунок того, що ринок готовий прийняти більш високу ціну. Це саме те, що нам і потрібно.

4.4.3 Вибір стратегії конкурентної поведінки

Наступним кроком виберемо стратегії конкурентної поведінки (табл. 4.16).

Таблиця 4.16 - Визначення базової стратегії конкурентної поведінки

№ п/п	Чи є проект “першопрохідцем” на ринку?	Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?	Чи буде компанія копіювати основні характеристики товару конкурента, і які?	Стратегія конкурентної поведінки
1	На локальному так. На міжнародному є схожі проекти	Забирати існуючих у держави та шукати нових	Частково. Буде здійснено покращення існуючих методів	Стратегія лідера

При визначенні стратегії конкурентної поведінки ми дійшли висновку, що нам доцільніше використовувати стратегію лідера, оскільки усі інші стратегії нам не підходять. Проект не оцінюється як другий чи третій на ринку, і не буде наслідувати державну систему реєстрації та підтвердження авторських прав.

4.4.4 Визначення стратегії позиціонування

На основі вимог споживачів з обраних сегментів до стартап-компанії та до продукту (див. табл. 4.5), а також в залежності від обраної базової стратегії розвитку (див. табл. 4.15) та стратегії конкурентної поведінки (див. табл. 4.16) розробимо стратегію позиціонування, що полягає у формуванні ринкової позиції (комплексу асоціацій), за яким споживачі мають ідентифікувати проект (табл. 4.17).

Таблиця 4.17 - Визначення стратегії позиціонування

№ п/п	Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкуренто-спроможні позиції власного стартап-проекту	Вибір асоціацій, які мають сформувати комплексну позицію власного проекту (три ключових)
1	Висока якість, швидкість надання послуг, співвідношення якість / вартість	Стратегія диференціації	Швидкість надання послуг з підтвердження майнових і авторських прав, метод хешування творів для оптимального зберігання та виявлення дублікатів чи трохи змінених оригіналів	Якість, швидкість, надійність

Головними вимогами цільовою аудиторії є висока якість та швидкість надання послуг, співвідношення ціни та якості послуг. Для забезпечення цього було обрано стратегію диференціації, а ключовими конкурентоспроможними позиціями стартап-проекту є використання нового алгоритму, який дозволить оптимально зберігати твори у вигляді хешів, що нададуть можливість виявлення дублікатів чи трохи змінених оригіналів. Ключові асоціації, які мають сформувати комплексну позицію власного проекту – якість, швидкість та надійність.

4.5 Розроблення маркетингової програми стартап-проекту

4.5.1 Формування маркетингової концепції товару

Сформуємо маркетингову концепцію товару, який отримає споживач. У таблиці 4.18 наведено результати попереднього аналізу конкурентоспроможності товару.

Таблиця 4.18 - Визначення ключових переваг концепції потенційного товару

№ п/п	Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами (існуючі або такі, що потрібно створити)
1	2	3	4
1	Підвищення рівня захисту	Захист відбувається за ліцензією держави, твір неможливо дістати із самої системи	Метод хешування вмісту творів, що дозволить виявляти дублікати та схожі за змістом
2	Легке підтвердження прав	Просте і швидке віднайдження всього ланцюга передачі прав від моменту реєстрації	Зберігання та захист від підробки всіх транзакцій починаючи від реєстрації твору до поточного власника, враховуючи також часткові передавання прав

Продовження таблиці 4.18

1	2	3	4
3	Швидка реєстрація нових творів	Швидкий та зручний інтерфейс, швидка перевірка наданих даних та документів	Реєстрація творів онлайн з будь-якого місця, оплата послуг онлайн, перевірка даних та документів під час подачі заявки онлайн
4	Швидка реєстрація передавання майнових прав на твір	Швидкий та зручний інтерфейс, швидка перевірка наданих даних та документів	Реєстрація повної або часткової передачі майнових на твори онлайн з будь-якого місця, оплата послуг онлайн, перевірка даних та документів під час подачі заявки онлайн

Формування маркетингової концепції проекту показало, що ключовими його перевагами є підвищений рівень захисту прав, швидка та зручна реєстрація нових творів та передача прав на вже існуючі, швидке та легке підтвердження прав поточним власником, не зважаючи на те, скільки ланок складає ланцюг передачі прав з часом.

4.5.2 Трирівнева маркетингова модель послуг

Результат розробки трирівневої маркетингової моделі товару наведено у таблиці 4.19.

Таблиця 4.19 - Опис трьох рівнів моделі товару

Рівні товару	Сутність та складові
1	2
I. Товар за задумом	Інформаційна система реєстрації та підтвердження авторських прав

Продовження таблиці 4.19

1	2		
II. Товар у реальному виконанні	Властивості/характеристики	М/Нм	Вр/Тх /Тл/Е/Ор
	1. Швидка реєстрація 2. Швидка передача власності 3. Швидке підтвердження власності чи авторства		
	Якість: відповідність вимогам SaaS		
	Пакування: хостинг у “хмарі” та доставка через Інтернет		
	Марка: Copyright registration and confirmation system (CRCS)		
III. Товар із підкріпленням	До продажу: не потребує особливих вимог		
	Після продажу: постійна цілодобова підтримка клієнтів		
За рахунок чого потенційний товар буде захищено від копіювання: за рахунок комерційної таємниці			

Основний задум даного стартап-проекту полягає у реалізації швидких та зручних послуг, надання доступу до системи через інтернет у будь-якому місці та у будь-яку пору доби. Захист системи від копіювання планується забезпечувати за рахунок збереження комерційної таємниці.

4.5.3 Визначення меж встановлення ціни

Визначення цінових меж, якими необхідно керуватись при встановленні ціни на потенційний товар, наведено у таблиці 4.20. Межі встановлено виходячи з аналізу рівня доходів цільової групи споживачів.

Таблиця 4.20 - Визначення меж встановлення ціни

№ п/п	Рівень цін на товари-замінники	Рівень цін на товари-аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на товар/послугу
1	від 63,75 (реєстрація для фізичних осіб) до декількох тисяч грн (залежить від розцінок юридичних компаній)	від 63,75 (реєстрація для фізичних осіб) до декількох тисяч грн (залежить від розцінок юридичних компаній)	від 7500 до 100000 грн (приблизно)	від 200 до 1000 грн (попередньо) для фізичних осіб договірна - для юридичних

4.5.4 Формування системи збуту

Сьогодні пошук товарів частіше за все відбувається через інтернет. Оскільки ми можемо продавати послуги безпосередньо клієнту, а також за допомогою посередників (партнерів), то глибина каналу збуту буде від нульового до першого рівня, про що свідчать дані таблиці 4.21.

Таблиця 4.21 - Формування системи збуту

№ п/п	Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
1	Купівля послуги відбувається під час замовлення послуги, або заздалегідь з укладанням контрактів	пошук потенційних клієнтів аналіз ринку програми лояльності реклама	Канал нульового рівня та однорівневий	Через Інтернет

4.5.5 Концепція маркетингових комунікацій

Останньою складовою маркетингової програми є розроблення концепції маркетингових комунікацій, що спирається на попередньо обрану основу для позиціонування, визначену специфіку поведінки клієнтів. Результати розробки даної концепції наведено у таблиці 4.22.

Таблиця 4.22 - Концепція маркетингових комунікацій

№ п/п	Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламного повідомлення	Концепція рекламного звернення
1	Клієнти дізнаються про новинки з інтернету, та шукають альтернативні державним послуги з кращими умовами та якістю	Соціальні мережі, статті, інтернет	Інтернет-маркетинг, маркетинг у соціальних мережах, реклама на телебаченні, презентації на конференціях	Донести до потенційних клієнтів альтернативу державним послугам із захисту авторських прав	Авторське право чи майнове право на твір клієнта під надійним захистом і довести його не потребує значних зусиль, як раніше

Оскільки клієнти більшу частину інформації про товари та послуги дізнаються через такі комунікації як: інтернет, соціальні мережі, телевізійна реклама, в якості ключових позицій для позиціонування було обрано інтернет-маркетинг, маркетинг у соціальних мережах, реклама на телебаченні. Також через специфіку сфери захисту авторських прав та її цільової аудиторії було додано до ключових позицій ще й презентації на конференціях.

4.6 Висновки по розділу

Проаналізувавши сильні та слабкі сторони ідеї, її можливості та загрози, попит, динаміку ринку та рентабельність проекту можна зробити висновок, що наша ідея є затребуваною, причому потенціальний попит на неї з моменту запуску проекту за прогнозами має тільки рости. Основним і єдиним конкурентом на внутрішньому ринку являється лише державна система захисту авторських прав, вирішенням проблем якої наша система має значні переваги та привабливість для потенційних клієнтів. Завдяки унікальним властивостям системи реєстрації та підтвердження авторських прав, у нас є дуже велика ймовірність зайняти лідерську позицію у сегменті захисту авторських прав на території України. Подальша імплементація проекту є доцільною, причому повернення інвестицій потенційно може відбутися вже на початку другого року після запуску системи.

ВИСНОВКИ

У даній роботі досліджено існуючі процеси реєстрації та підтвердження авторських прав в системі державного захисту авторських прав в Україні. У ході виконання аналізу покращено усі процеси за рахунок автоматизації їх основних частин, а особливо процес підтвердження права на твір з віднайденням всього ланцюга часткової та повної передачі прав від реєстрації до поточного власника (або власників). Розробка нової системи реєстрації та підтвердження авторських прав підвищить рівень їх захисту та імплементує у собі рішення проблем доведення авторських прав.

У даній роботі також проаналізовано та вдосконалено метод отримання унікальних цифрових відбитків документів у вигляді хешів, які створюють колізію у разі хешування цим методом подібних документів. Це дозволить розробити систему, яка надаватиме можливості первинної реєстрації творів авторами, реєстрації передачі авторських прав між фізичними та юридичними особами та підтвердження (доведення) авторських прав чи прав власності на твори, однією із головних особливостей якої буде збереження цифрового відбитку твору замість збереження оригінального контенту файлу у вигляді хешу, що дозволить виявити плагіат чи близький за змістом твору разі повторної спроби реєстрації такого твору в системі.

Систему реєстрації та підтвердження авторських прав можливо розробити за допомогою відкритих та безкоштовних засобів розробки та управління проектами. Розгортання системи має класичну клієнт-серверну архітектуру та доставляється клієнтам як веб-додаток засобами мережі Інтернет. Початковий вигляд архітектури системи не є складним, тому для її імплементації не потрібно наймати спеціалістів найвищого класу, що спростить пошук необхідних кадрів та зменшить витрати на їх утримання.

Проаналізувавши сильні та слабкі сторони ідеї, її можливості та загрози, попит, динаміку ринку та рентабельність проекту можна зробити висновок, що наша

ідея є затребуваною, причому потенціальний попит на неї з моменту запуску проекту за прогнозами має тільки рости. Основним і єдиним конкурентом на внутрішньому ринку являється лише державна система захисту авторських прав, вирішенням проблем якої наша система має значні переваги та привабливість для потенційних клієнтів. Завдяки унікальним властивостям системи реєстрації та підтвердження авторських прав, у нас є дуже велика ймовірність зайняти лідерську позицію у сегменті захисту авторських прав на території України. Подальша імплементація проекту є доцільною, причому повернення інвестицій потенційно може відбутися вже на початку другого року після запуску системи.

Отже, враховуючи усі висновки, що ми отримали у ході виконання даної роботи, вважатимемо мету даної роботи досягнутою.

ПЕРЕЛІК ПОСИЛАНЬ

[1] Сіліна М. Особливості процесу доказування під час судового розгляду справ про порушення авторського права. Режим доступу: <http://llg.com.ua/osoblivosti-procesu-dokazuvannya-pid-chas-sudovogo-rozglyadu-sprav-pro-porushennya-avtorskogo-prava/>.

[2] Реєстрація авторських прав в Україні — алгоритм. Компанія “Wise”. Режим доступу: <https://wisegroup.com.ua/ua/reestraciya-avtorskix-prav-v-ukra%D1%97ni-algorithm/>.

[3] Реєстрація авторських прав. Компанія “Inventa”. Режим доступу: <https://www.inventa.ua/ua/nashi-poslugi/reestratsiya-avtorskih-prav/>.

[4] Що таке авторське право? РІК, 2018. Режим доступу: <https://pik.org.ua/shho-take-avtorske-pravo.html>.

[5] Захист авторських прав в інтернеті: нові техніки плагіату. Checkpoint Business Media, 2016. Режим доступу: <https://ckp.in.ua/articles/11504>.

[6] Веб-технології для розробників. JavaScript. Mozilla Development Network, 2019. Режим доступу: <https://developer.mozilla.org/uk/docs/Web/JavaScript>.

[7] Xiaoyun Wang and Dengguo Feng and Xuejia Lai and Hongbo Yu (received 16 Aug 2004, last revised 17 Aug 2004). Collisions for Hash Functions MD4, MD5, HAVAL-128 and RIPEMD. Cryptology ePrint Archive: Report 2004/199.

[8] Sanadhya S.K., Sarkar P. (2008) Deterministic Constructions of 21-Step Collisions for the SHA-2 Hash Family. In: Wu TC., Lei CL., Rijmen V., Lee DT. (eds) Information Security. ISC 2008. Lecture Notes in Computer Science, vol 5222. Springer, Berlin, Heidelberg.

[9] Morris J. Dworkin. SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions. Federal Inf. Process. Stds. (NIST FIPS) - 202, 2015.

[10] U. Manber. Finding Similar Files in a Large File System. Winter USENIX Technical Conference, 1994.

[11] A. Shrivastava. Exact Weighted Minwise Hashing in Constant Time. arXiv preprint arXiv:1602.08393, 2016.

[12] Moses S. Charikar. Similarity estimation techniques from rounding algorithms. Proceedings of the 34th Annual ACM Symposium on Theory of Computing, 2002, p. 380.

[13] Sadowski, C., & Levin, G. (2007). Simhash: Hash-based similarity detection. Citeseer. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.7179&rep=rep1&type=pdf>.

[14] Терещенко Р. В. Використання технології блокчейн для захисту авторських прав // Чотирнадцята міжнародна науково-практична конференція “Математичне та імітаційне моделювання систем. МОДС '2019”. Матеріали конференції. – Чернігів – 24-26 червня 2019 р.

[15] Терещенко Р. В., Жданова О.Г. Система реєстрації та підтвердження авторських прав // Тези III всеукраїнської науково-практичної конференції молодих вчених та студентів “Інформаційні системи та технології управління” (ІСТУ-2019) – м. Київ.: НТУУ “КПІ ім. Ігоря Сікорського”, 20-22 листопада 2019 р. – с. 205.

[16] Брюс Шнайер. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. — М.: Триумф, 2002. — 816 с. — 3000 экз.

[17] Лапони́на О.Р. Криптографические основы безопасности. — М.: Интернет-университет информационных технологий - ИНТУИТ.ру, 2004. — 320 с.

[18] Wu H. The Hash Function JH. // <http://csrc.nist.gov> – January 15, 2009 – Institute for Infocomm Research, Singapore.

[19] Bertoni G., Daemen J., Peeters M., Van Assche G. Keccak sponge function family main document. // <http://keccak.noekeon.org> – Version 1.2 – 23.04.2009.

[20] А. Щербаков, А. Домашев. Прикладная криптография. — М.: Русская Редакция, 2003. — 404 с.

[21] N. Ferguson, S. Lucks, B. Schneier. The Skein Hash Function Family. // <https://www.schneier.com/academic/skein/> – Version 1.3 — 1 Oct 2010.

[22] Eli Biham, Adi Shamir. Differential cryptanalysis of Snefru, Khafre, REDOC-II, LOKI and Lucifer (Extended Abstract).

[23] John Kelsey, Stefan Lucks. Collisions and Near-Collisions for Reduced-Round Tiger // Graz, 2006.

[24] Rivest R. The MD5 Message-Digest Algorithm. // <http://www.ietf.org> – April 1992.

[25] Панасенко С.П. Алгоритм хэширования MD4: обзор криптоаналитических исследований. // <http://daily.sec.ru> – 2012.

[26] Menezes A., van Oorschot P., Vanstone S. Handbook of Applied Cryptography. - CRC Press, 1996 // <http://www.cacr.math.uwaterloo.ca>.

[27] Y.Zheng, J. Pieprzyk, and J. Seberry. HAVAL A One-Way Hashing Algorithm with Variable Length of Output, Advances in Cryptology AUSCRYPT 1992 Proceedings, Springer-Verlag, 1993, p. 83-104.

[28] B. den Boer and A. Bosselaers. Collisions for the Compression Function of MD5, Advances in Cryptology EUROCRYPT 93 Proceedings, Springer-Verlag, 1994, pp. 293.

[29] Кудін А., Коваленко Б. Алгоритмічні аспекти пошуку прообразів геш-функцій на прикладі MD5. / Захист інформації, том 17, № 3, липень-вересень 2015.

[30] Глинчук Л.Я. Криптологія. Навчально-методичний посібник. / Луцьк, Вежа-Друк, 2014 р - 186 с.

[31] B. Kaliski. The MD2 Message-Digest Algorithm. / RSA Laboratories, April 1992. Режим доступу: <https://tools.ietf.org/html/rfc1319>.

[32] R. Rivest. The MD4 Message Digest Algorithm. / MIT Laboratory for Computer Science, October 1990. Режим доступу: <https://tools.ietf.org/html/rfc1186>.

[33] R. Rivest. The MD5 Message-Digest Algorithm. / MIT Laboratory for Computer Science and RSA Data Security, Inc., April 1992. Режим доступу: <https://tools.ietf.org/html/rfc1321>.

[34] Ronald L. Rivest. The MD6 hash function. A proposal to NIST for SHA-3. / Massachusetts Institute of Technology Cambridge, October 2008.

[35] Y. Zheng, J. Pieprzyk and J. Seberry. HAVAL — a one-way hashing algorithm with variable length of output. / Advances in Cryptology — AUSCRYPT'92, Lecture Notes in Computer Science, Vol.718, pp. 83—104, Springer-Verlag, 1993.

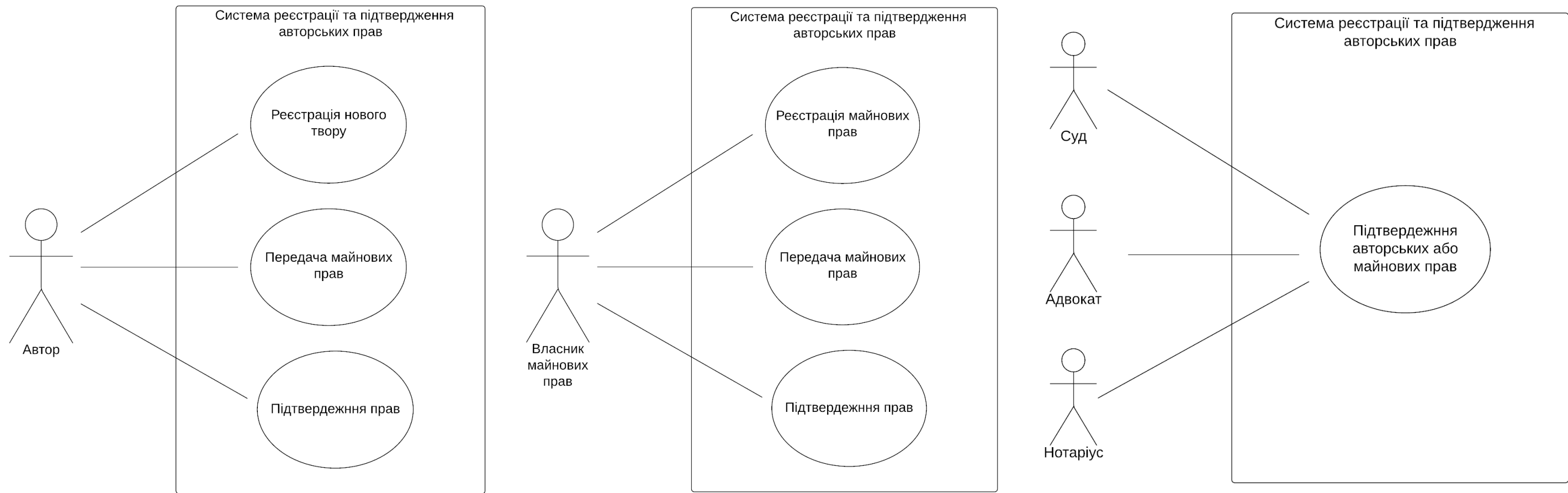
[36] D. Eastlake, P. Jones. US Secure Hash Algorithm 1 (SHA1). / Motorola, Cisco Systems, September 2001. Режим доступа: <https://tools.ietf.org/html/rfc3174>.

[37] Proposed ECMAScript 4th Edition – Language Overview. / Adobe Systems Inc., The Mozilla Foundation, Opera Software ASA, and others, October 2007. Режим доступа: <https://www.webcitation.org/5rBiWD4P6?url=http://www.ecmascript.org/es4/spec/overview.pdf>.

ДОДАТОК А

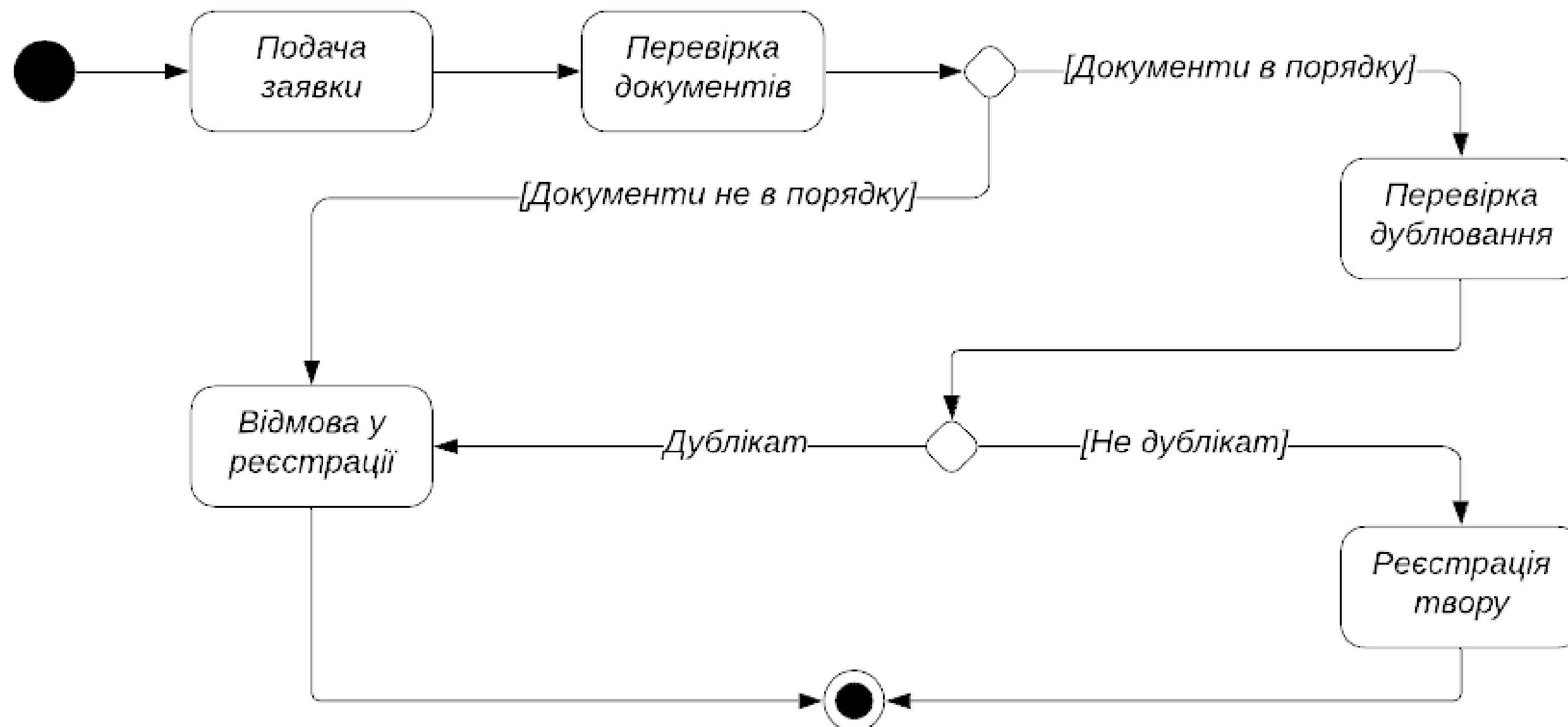
Графічний матеріал

Діаграми способів використання системи для автора, власника майнових прав та інших акторів



Демонстраційний плакат до магістерської дисертації
“СИСТЕМА РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ”
Магістрант
Керівник
Терещенко Р.В.
Жданова О.Г.

Діаграма процесу первинної реєстрації твору автором



Демонстраційний плакат до магістерської дисертації
 “СИСТЕМА РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ”

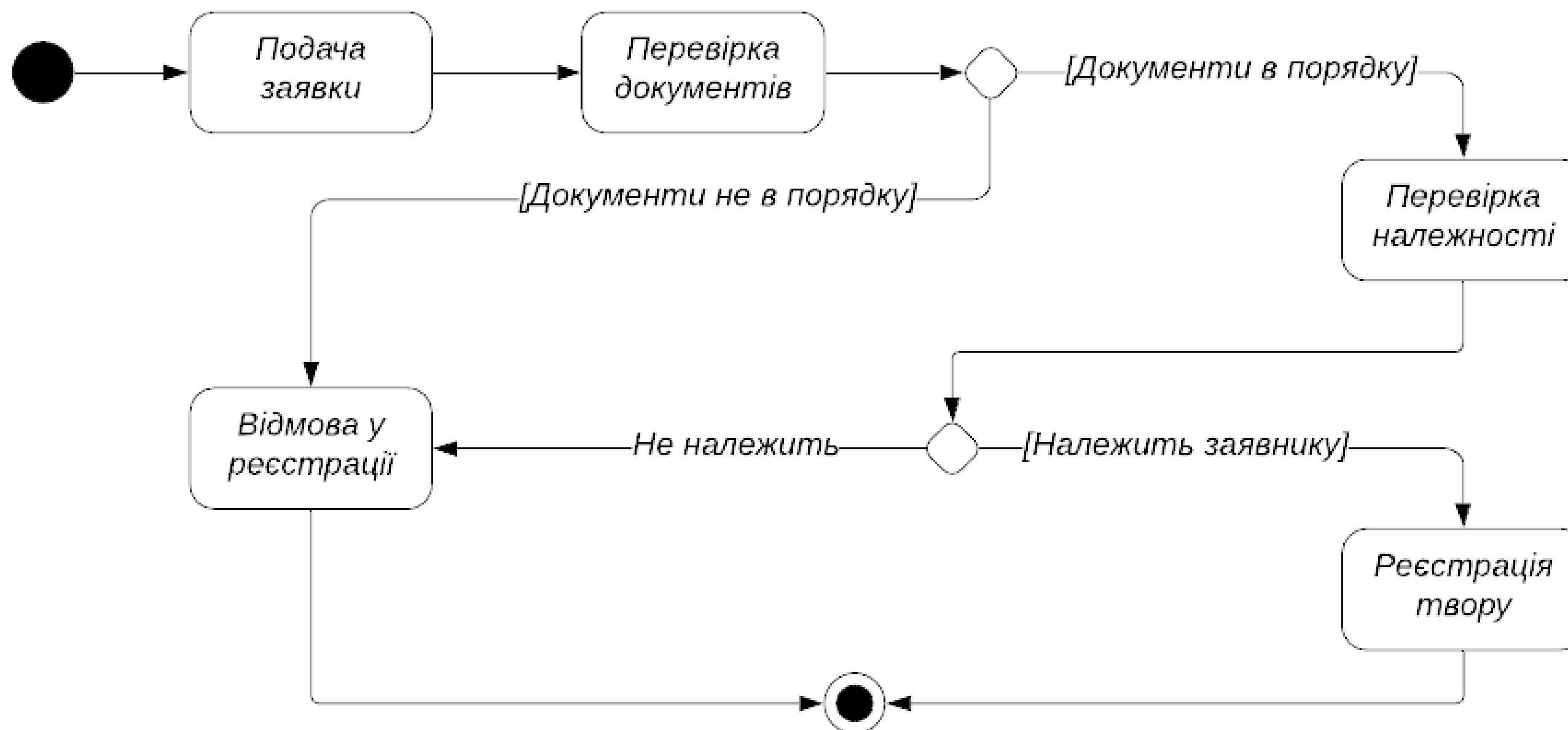
Магістрант

Терещенко Р.В.

Керівник

Жданова О.Г.

Діаграма процесу передачі майнових прав на твір



Демонстраційний плакат до магістерської дисертації
 “СИСТЕМА РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ”

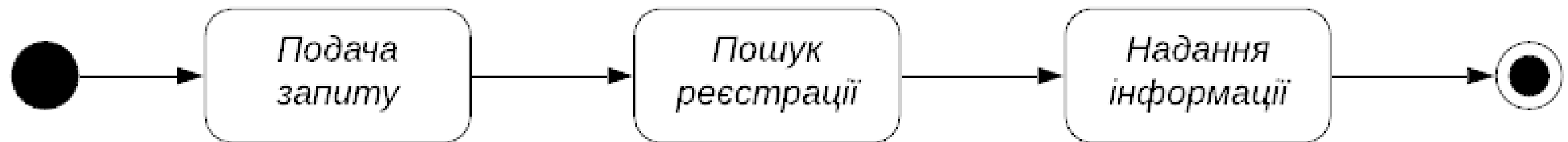
Магістрант

Терещенко Р.В.

Керівник

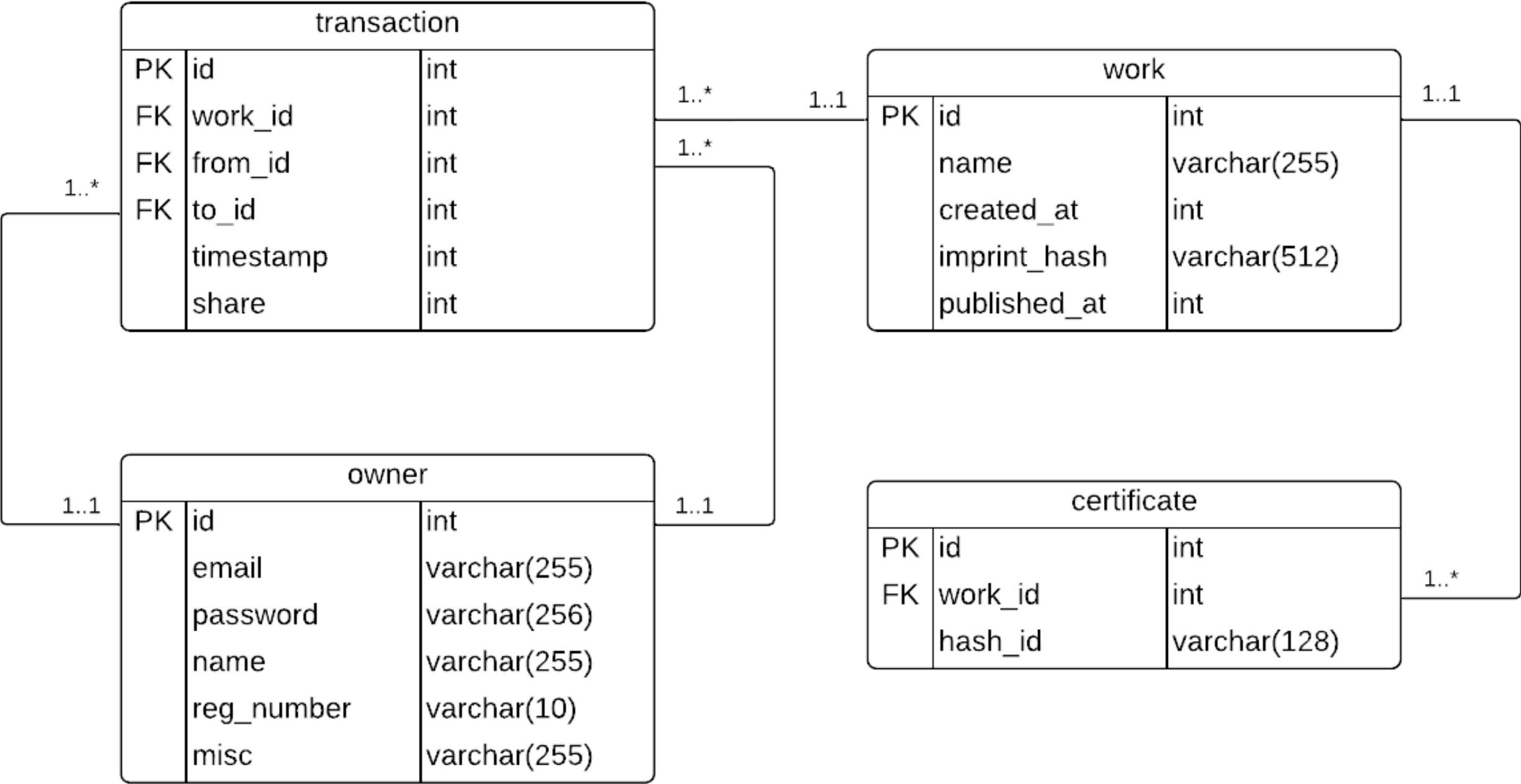
Жданова О.Г.

Діаграма процесу підтвердження авторських та майнових прав на твір



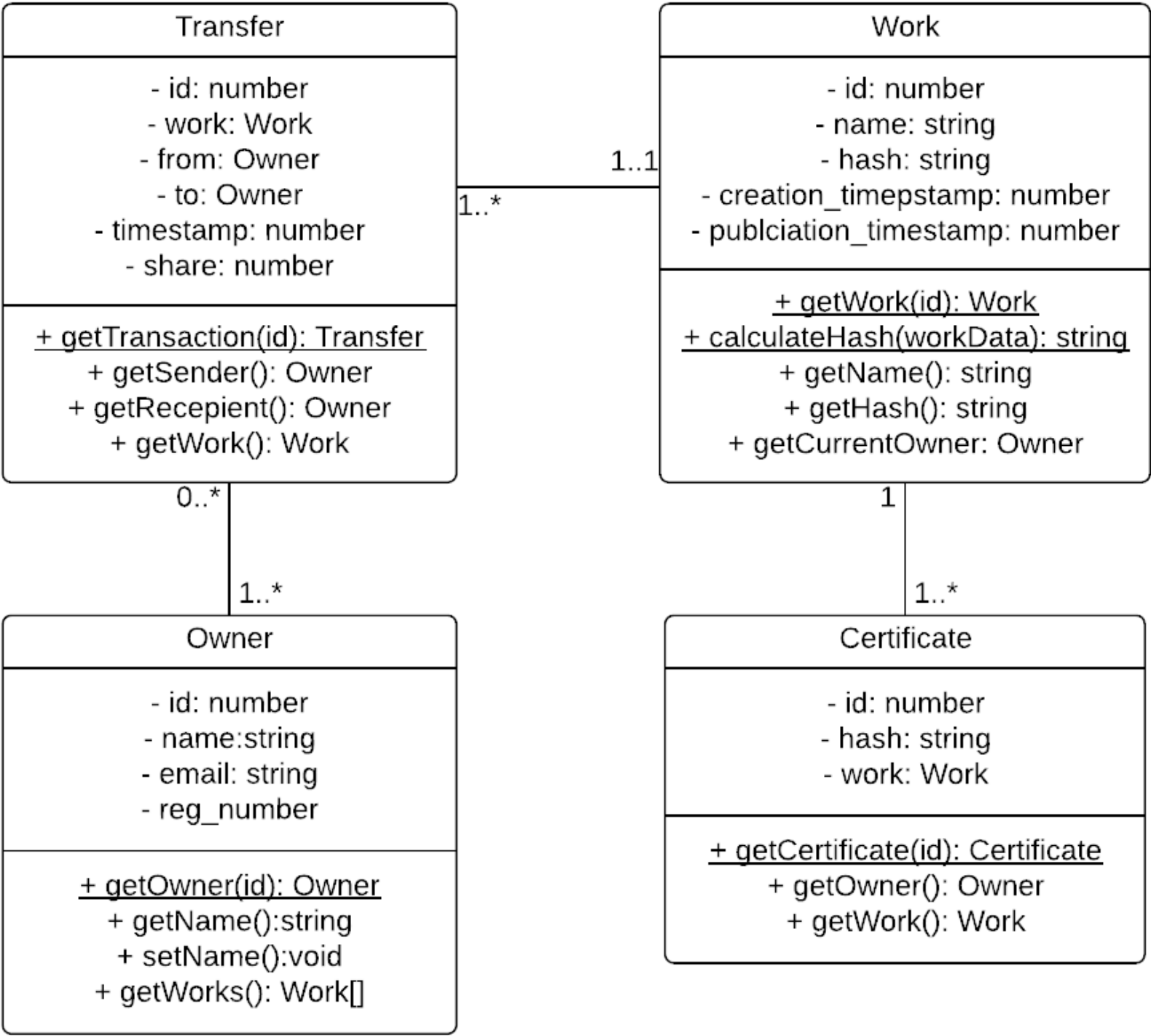
Демонстраційний плакат до магістерської дисертації
“СИСТЕМА РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ”
Магістрант
Керівник
Терещенко Р.В.
Жданова О.Г.

Діаграма інформаційної моделі системи реєстрації та підтвердження авторських прав



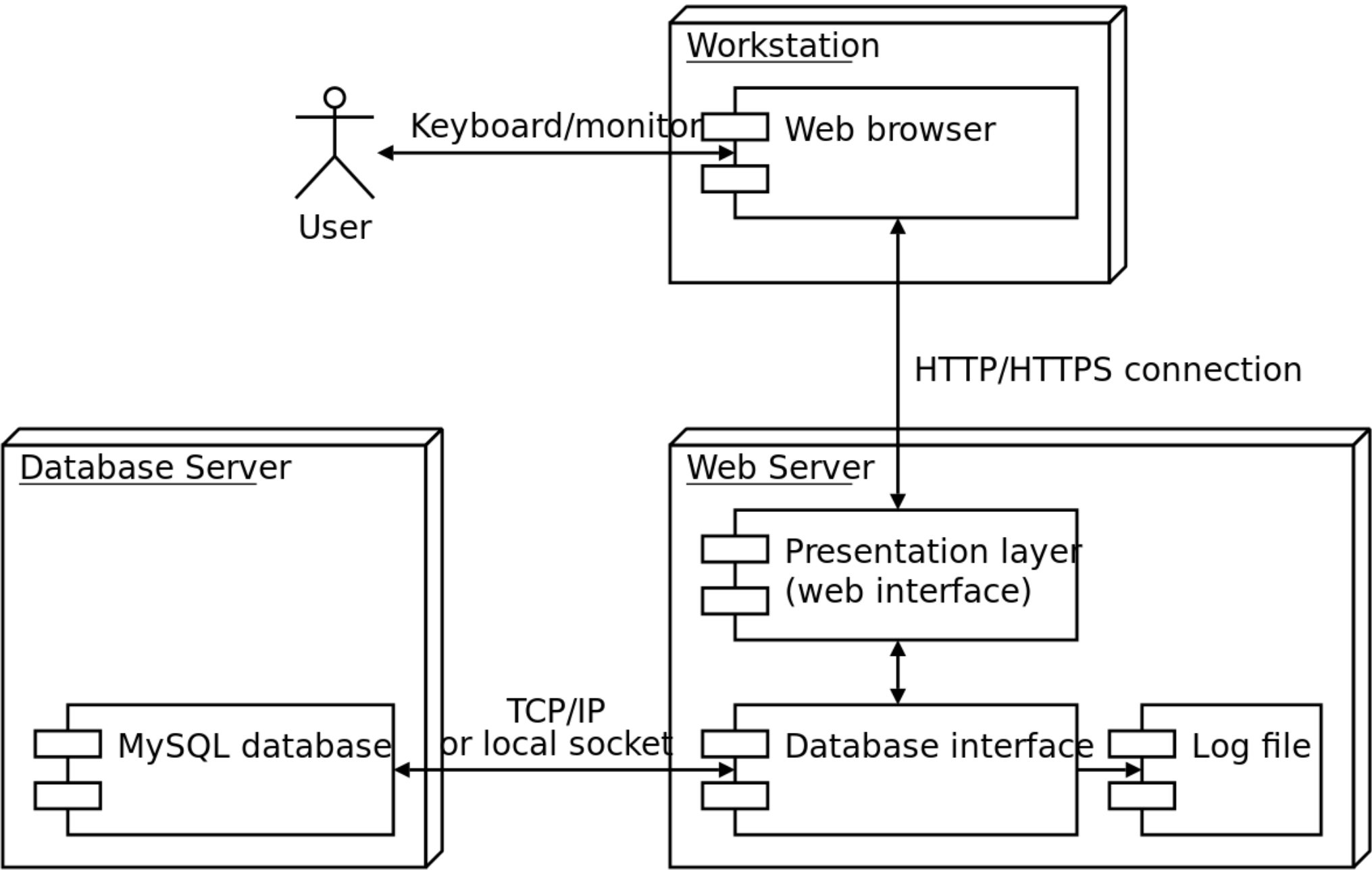
Демонстраційний плакат до магістерської дисертації
“СИСТЕМА РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ”
Магістрант
Керівник
Терещенко Р.В.
Жданова О.Г.

Діаграма класів інформаційної системи реєстрації та підтвердження авторських прав



Демонстраційний плакат до магістерської дисертації
“СИСТЕМА РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ”
Магістрант
Керівник
Терещенко Р.В.
Жданова О.Г.

Діаграма розгортання інформаційної системи реєстрації та підтвердження авторських прав



Демонстраційний плакат до магістерської дисертації
“СИСТЕМА РЕЄСТРАЦІЇ ТА ПІДТВЕРДЖЕННЯ АВТОРСЬКИХ ПРАВ”
Магістрант
Керівник
Терещенко Р.В.
Жданова О.Г.